Dietary Fiber and Colorectal Cancer Risk: A Nested Case–Control Study Using Food Diaries


Manuscript received June 6, 2009; revised February 1, 2010; accepted February 26, 2010.
† Died June 16, 2009, between submission and acceptance of this article.
Correspondence to: Kay Tee Khaw, Medical Research Council Centre for Nutritional Epidemiology in Cancer Prevention and Survival, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK (e-mail: kk101@medschl.cam.ac.uk).

Background
Results of epidemiological studies of dietary fiber and colorectal cancer risk have not been consistent, possibly because of attenuation of associations due to measurement error in dietary exposure ascertainment.

Methods
To examine the association between dietary fiber intake and colorectal cancer risk, we conducted a prospective case–control study nested within seven UK cohort studies, which included 579 case patients who developed incident colorectal cancer and 1996 matched control subjects. We used standardized dietary data obtained from 4- to 7-day food diaries that were completed by all participants to calculate the odds ratios for colorectal, colon, and rectal cancers with the use of conditional logistic regression models that adjusted for relevant covariates. We also calculated odds ratios for colorectal cancer by using dietary data obtained from food-frequency questionnaires that were completed by most participants. All statistical tests were two-sided.

Results
Intakes of absolute fiber and of fiber intake density, ascertained by food diaries, were statistically significantly inversely associated with the risks of colorectal and colon cancers in both age-adjusted models and multivariable models that adjusted for age; anthropomorphic and socioeconomic factors; and dietary intakes of folate, alcohol, and energy. For example, the multivariable-adjusted odds ratio of colorectal cancer for highest vs lowest quintile of fiber intake density was 0.66 (95% confidence interval = 0.45 to 0.96). However, no statistically significant association was observed when the same analysis was conducted using dietary data obtained by food-frequency questionnaire (multivariable odds ratio = 0.88, 95% confidence interval = 0.57 to 1.36).

Conclusions
Intake of dietary fiber is inversely associated with colorectal cancer risk. Methodological differences (ie, study design, dietary assessment instruments, definition of fiber) may account for the lack of convincing evidence for the inverse association between fiber intake and colorectal cancer risk in some previous studies.


Environmental and lifestyle factors are believed to play a large role in the incidence of colorectal cancer (1), which is currently the third most common cancer in the world with more than a million incident cases estimated in 2002 (2). The fact that known high-penetrance gene variants that are associated with colorectal cancer risk explain fewer than 5% of the observed cases (3), together with the wide geographical differences in colorectal cancer risk and the marked secular changes in colorectal cancer rates within certain populations and in studies of migrants (4–7), suggests that environmental and lifestyle factors including diet are important factors that influence risk. Of the many environmental and dietary factors thought to be involved in risk modification, dietary fiber has long been thought to be associated with a reduced risk of colorectal cancer (8,9). Ecological international comparisons (10) and some case–control studies (11,12) support this association, and well-established mechanisms have been identified in human and animal experimental systems whereby fiber entering the colonic lumen inhibits carcinogenesis within the colorectal mucosa (13,14).

Nevertheless, analytic epidemiological studies of dietary fiber and the risk of colorectal cancer have not yielded consistent associations (15–23), and results of the World Cancer Research Fund and American Institute for Cancer Research (WCRF/AICR) meta-analysis were deemed to have yielded “probable” rather than “convincing” evidence of an association (24). For example, in the European Prospective Investigation into Cancer and Nutrition (EPIC), a large prospective study with 1065 colorectal cancer case patients in a study population of 520000 individuals throughout Europe, the risk of colorectal cancer was reduced by 40% in the highest vs lowest quintile of fiber intake (relative risk [RR] = 0.58, 95% confidence interval [CI] = 0.41 to 0.85) (15). In the National Institutes of Health–AARP Diet and Health Study (16), which evaluated 2974 colorectal cancer case patients in a cohort of
Articles based on pooled standardized data from 579 case patients with cancer among those who consumed less than 10 g of fiber per day compared with those who consumed 10–15 g of fiber per day (RR = 1.18, 95% CI = 1.05 to 1.31) (17). A study of 37,562 individuals who were screened for colorectal adenoma in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial observed an inverse association between fiber intake and colorectal cancer risk (RR for highest vs lowest quintile of intake = 0.73, 95% CI = 0.62 to 0.86) (18). Intervention studies of the effect of dietary and supplemental fiber on colorectal adenoma recurrence have shown null (19–21), or even adverse (22), effects. However, follow-up time in the intervention studies was, in general, short, and a pooled reanalysis of two of these trials showed a statistically significant interaction by sex and a beneficial effect of the intervention in men (odds ratio [OR] = 0.81, 95% CI = 0.67 to 0.98) (23).

Measuring a subject’s usual intake of dietary fiber and other food constituents in these and other large prospective dietary studies is difficult. Data on food portions and types are almost universally assessed by food-frequency questionnaires (FFQs). However, data from FFQs harbor systematic errors according to participants’ age, sex, and body mass index, as well as random errors (25–27), and the ability of FFQs to assess more than very large underlying associations between diet and cancer has been questioned (26,28,29).

Since 1982, several prospective cohorts in the United Kingdom have used additional methods to assess diet. In an initial pilot study comparing several epidemiological dietary assessment methods, intakes estimated from prospective food diaries completed by subjects showed higher correlations with biomarkers of dietary intake, such as urinary nitrogen and potassium, than data from an FFQ (26,30). Consequently, most of the UK cohorts have measured diet by using both an FFQ and 4- to 7-day food diaries. However, because analysis of food diaries is resource intensive, few cohorts have reported results using these methods. Nevertheless, two nested case–control studies of fat intake and breast cancer risk—one from the United Kingdom (31) and the other from the United States (32)—reported good evidence of an association between dietary fat intake and breast cancer risk using dietary data derived from food diaries but not when using data derived from FFQs that were completed by the same participants.

The UK cohorts are now mature, and in 2006, funds for the analysis of these food diaries using a standardized data entry system for case patients diagnosed with colorectal, breast, or prostate cancer and matched control subjects were awarded under the auspices of the Medical Research Council Centre for Nutritional Epidemiology in Cancer Protection and Survival (CNC). The UK cohort consortium comprises seven cohorts with a total cohort size of 153,000 individuals (www.srl.cam.ac.uk/cnc). Here, we report the association between dietary fiber and colorectal cancer risk based on pooled standardized data from 579 case patients with incident colorectal cancer and 1996 matched control subjects. We used a standardized method of assessing the content of dietary fiber in food, which is expressed as nonstarch polysaccharides (NSP) (33) in the UK food composition tables (34), to assess individual total dietary fiber intake.

Participants and Methods

Participants

The dietary data were ascertained via food diaries that were completed at study recruitment and/or a subsequent monitoring phase (baseline) by participants in seven established UK cohorts: EPIC-Norfolk, EPIC-Oxford, the Guernsey Study, the Medical Research Council National Survey of Health and Development (NSHD), the Oxford Vegetarian Study, the UK Women’s Cohort Study (UKWCS), and Whitehall II (Table 1). The methods of recruitment, study design, and ethical approval for each of these cohorts have been described in detail elsewhere (35–40). Each cohort collected dietary information with the use of 4-day (Guernsey Study, Oxford Vegetarian Study, and UKWCS) (36,38), 5-day (NSHD) (40), or 7-day (EPIC-Norfolk, EPIC-Oxford, and Whitehall II) (35,37,39) food diaries that were completed on consecutive days. Participants were asked to record in detail all the foods and beverages they consumed at designated times throughout the day, usually
as prompted by time slots such as “mid-morning—between break- 
fast time and lunchtime” and by photographs of standard-sized 
plates with three different portion sizes of representative foods to 
help participants estimate the amounts they consumed (41,42). 
Information on age, sex, height, weight, smoking habits, education 
level, socioeconomic status, physical activity, family history of co-
lorectal cancer, and use of aspirin were collected either in inter-
views conducted by trained researchers or via questionnaires that 
were administered before completion of the food diary. FFQs 
comprising 127–217 items, which were based on the FFQ that was 
used in the US Nurses’ Health Study and validated for use in the 
United Kingdom (36,43,44), were also administered before com-
pletion of the food diaries and were available for analysis for most 
of the participants in the EPIC-Norfolk, EPIC-Oxford, UKWCS, 
and Whitehall II cohorts.

Follow-up and Ascertainment of Colorectal Cancer Case 
Patients
Case patients were individuals who reported that they had not 
been diagnosed with any nonmelanoma cancer at baseline and who 
were diagnosed with colorectal cancer more than 12 months 
after the date they began completing the food diary but before the 
end of the study period, which was defined for each study center 
as the latest date of complete follow-up for both cancer incidence 
and vital status. Follow-up for diagnosis of colorectal cancer was 
provided through record linkage with the UK Office of National 
Statistics and local cancer registries (eg, the Eastern Cancer 
Registration and Information Center). Cases of colorectal cancer 
were defined using codes C18–C20 from the Tenth Revision of the 
International Statistical Classification of Diseases and Related Health 
Problems (45). The last dates of follow-up varied among the 
cohorts and ranged from December 31, 2003, to January 1, 2007 
(Table 1).

Selection of Matched Control Subjects
Case patients were matched within their respective cohort to four 
control subjects each with the following exceptions: Some case 
patients from EPIC-Oxford, the Guernsey Study, and the Oxford 
Vegetarian Study were matched to two control subjects, and some 
from the UKWCS were matched to five control subjects. Matched 
control subjects were selected at random from the appropriate 
stratum of the set of all cohort members who were alive at the end 
of follow-up. Matching criteria were sex, age at enrollment (±3 
years), and date of diary completion (±3 months). Control subjects 
who reported a prevalent nonmelanoma cancer diagnosis at base-
line were excluded.

Food Diary Coding
The majority of the food diary data were coded to give nutrient 
intakes and food group information by using the CNC data entry 
and processing programs Data Into Nutrients for Epidemiological 
Research (DINER) and DINERMO (46). A total of 51 of the 125 
UKWCS food diaries were coded and processed by using the Diet 
and Nutrition Tool for Evaluation (DANTE) (47) dietary assess-
ment program. We performed a comparative analysis of the 
DANTE and DINER/DINERMO programs: Selected nutrients 
from 100 randomly selected UKWCS food diaries were coded, 
checked, and calculated using both systems. There was good 
agreement for most nutrients calculated using the two data entry 
programs. However, the geometric mean intakes of energy and 
carbohydrate calculated using DANTE were 2% higher (95% 
CI = 0% to 5% higher) than those calculated using DINER. The 
geomeric mean intake of fiber calculated using DANTE was 8% 
higher (95% CI = 4% to 12% higher) than that calculated using 
DINER, which is equivalent to an arithmetic mean difference of 
1.3 g. Therefore, we treated UKWCS data calculated using 
DANTE separately from UKWCS data calculated using DINER

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Study aim</th>
<th>Year of diary completion</th>
<th>Last follow-up date</th>
<th>No. of case patients</th>
<th>No. of control subjects</th>
<th>Mean age at baseline, y (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPIC-Norfolk</td>
<td>Mortality and disease incidence in general population</td>
<td>1993–1998</td>
<td>December 31, 2006</td>
<td>318</td>
<td>1272</td>
<td>64.0 (7.9)</td>
</tr>
<tr>
<td>EPIC-Oxford</td>
<td>Mortality and disease incidence in general population and vegetarians</td>
<td>1993–1998</td>
<td>December 31, 2004</td>
<td>121</td>
<td>280</td>
<td>61.6 (10.6)</td>
</tr>
<tr>
<td>Guernsey Study</td>
<td>Cancer incidence in women living on Guernsey</td>
<td>1987–1991</td>
<td>December 31, 2003</td>
<td>28</td>
<td>55</td>
<td>59.3 (10.2)</td>
</tr>
<tr>
<td>MRC NSHD</td>
<td>Nationally representative birth cohort of men and women born in a week in March 1946 in England, Scotland, and Wales</td>
<td>1989</td>
<td>December 31, 2006</td>
<td>7</td>
<td>28</td>
<td>43 (0.0)</td>
</tr>
<tr>
<td>UKWCS</td>
<td>Mortality and cancer incidence in middle-aged women</td>
<td>1999–2003</td>
<td>December 31, 2006</td>
<td>25</td>
<td>100</td>
<td>63.1 (8.9)</td>
</tr>
</tbody>
</table>

* EPIC = European Prospective Investigation into Cancer and Nutrition; MRC NSHD = Medical Research Council National Survey of Health and Development; UKWCS = UK Women’s Cohort Study.
in all analyses. The 35 NSHD food diaries were coded using the Diet In Nutrients Out (DIDO) program (48). In a validation study using 100 NSHD food diaries, we found systematic differences in the measurements for many nutrients between the DIDO and DINNER/DINERMO programs (data not shown). However, although there were differences between these programs in fiber measurements within some food groups, overall, there was no statistically significant difference between the programs regarding measures of fiber intake. Measurements of energy intake were statistically significantly higher using DIDO compared with DINNER, and measurements of alcohol intake were statistically significantly lower. We believe that the differences between the measurements under the two systems reflect differences in portion sizes. We decided to retain the DIDO measurements for the NSHD food diaries because the portion sizes used in DIDO were contemporary to the dates of diary completion.

Statistical Analysis
Dietary fiber intake was expressed as absolute intake (grams per day), as intake density (grams per megajoule), and as the residuals from the regression of fiber intake on total energy intake. Sex-specific cut points for quintiles of dietary fiber intakes were derived from the distributions of dietary fiber intake, intake density, or residuals among control subjects. Conditional logistic regression models were used to calculate odds ratios of colorectal cancer and 95% confidence intervals. The odds ratio of colorectal cancer for a 6-g/d increase in dietary fiber intake density was calculated by assigning each increasing quintile a score of 1–5, respectively, and then using the score as a continuous variable in the regression model. Including a quadratic term for fiber intake did not improve the fit of our models. Visual inspection of histograms of food diary data identified a group of three outliers of fiber intake at values greater than 45 g/d and four outliers in fiber intake density at values greater than 5.5 g/MJ. The values for these outliers were truncated to the maximum value of the remaining data, that is, to 45 g/d and to 5.5 g/MJ for fiber intake and fiber intake density, respectively. Truncation of the values caused the SD of dietary fiber intake across the whole study population to shift from 5.84 to 5.65 g/d. We rounded the SD to the nearest whole number, 6 g/d, and calculated the odds ratio of colorectal cancer for a 6-g/d increase in dietary fiber intake. The SD of fiber intake density was 0.7 g/MJ before and after truncation; therefore, we calculated the odds ratio of colorectal cancer for a 0.7-g/MJ increase in fiber intake density.

For FFQ data, the distributions were truncated to 50 g/d (N = 5) and 5.5 g/MJ (N = 2) for fiber intake and fiber intake density, respectively, and the linear risk was calculated for the same unit increases as was done for the food diary data.

To examine the effects of potential confounders (other than the matching criteria, which were controlled for by design), the analyses were repeated by including the following variables in the conditional logistic regression models: height (meters, continuous), weight (kilograms, continuous), smoking status (never, past, current), physical activity (inactive, moderately inactive, moderately active, active), education level (no formal qualifications, lower secondary school to age 16 [UK: General Certificate of Secondary Education or equivalent], higher secondary school to age 18 [UK: A levels or equivalent], university degree), socioeconomic status [I = professional occupations; II = managerial and technical occupations; III-NM = skilled occupations, nonmanual; III-M = skilled occupations, manual; IV = partly skilled occupations; V = unskilled occupations (49)], alcohol intake (grams per day, continuous), dietary vitamin D intake (micrograms per day, continuous), dietary calcium intake (milligrams per day, continuous), and intakes of red and processed meats (grams per day, continuous). The categories for physical activity were defined as previously described (50) as follows: sedentary job and no recreational activity (inactive), sedentary job with less than 0.5 hour of recreational activity per day or standing job with no recreational activity (moderately inactive), sedentary job with 0.5–1 hours of recreational activity per day or standing job with less than 0.5 hour recreational activity per day or physical job with no recreational activity (moderately active), and sedentary job with more than 1 hour of recreational activity per day or standing job with more than 1 hour of recreational activity per day or physical job with at least some recreational activity or heavy manual job (active), or as similar to this index as possible based on individual cohorts’ data. Dietary fiber does not contribute directly to daily energy intake; however, it is more strongly associated with nonfat dietary components than with fat components (15). Therefore, we separated energy intake according to nutrient sources into intake from fat (megajoules per day, continuous; Pearson correlation coefficient with fiber intake = .20, P < .001) and nonfat (megajoules per day, continuous; Pearson correlation coefficient with fiber intake = .45, P < .001) sources to fully capture this information within the regression models for absolute fiber intake (15). It has been suggested that folate intake may confound findings on dietary intakes and colorectal cancer risk in European populations because folic acid fortification of foods is not mandatory in European countries (51) and in populations that have high usage of folate-containing multivitamins (52). Data on multivitamin use were not available for this study. However, a preliminary investigation indicated that general supplement use in the participating cohorts ranged from 30%, which is similar to the UK average, to 60% (data not shown). Therefore, we investigated the effect of adjustment for folate intake (micrograms per day; continuous).

Some cohorts did not collect information about some of the nondietary covariates. For example, information about physical activity is missing for all individuals in the Guernsey Study and the NSHD, and education level is missing for all individuals in the Oxford Vegetarian Study. Primary analyses (model a: unadjusted; model b: adjusted for height, weight, energy from fat and nonfat sources, and intakes of alcohol and folate) were restricted to the 2534 individuals for whom we had complete covariate information for height and weight. The analysis using model c, which additionally adjusted for smoking habits, socioeconomic status, physical activity, and education level, was necessarily restricted to those with complete covariate information; the numbers of case patients and control subjects included in this complete case analysis are given in table footnotes (Tables 3–5). Model d (presented in the text), which included further adjustment for dietary vitamin D intake, dietary calcium intake, and intakes of red and processed meats, was also based on this restricted set of participants. The main analyses were conducted with colorectal cancer as the outcome. We also investigated colon or rectal cancer outcomes. We conducted sensitivity analyses that excluded case patients who were diagnosed with colorectal cancer within the first 5 years of follow-up, to test the robustness of the main findings.
cancer within 2 years of diary completion. Because matching of case patients to control subjects included age (±3 years), we assessed the possibility of residual confounding by age by adjusting for age in years in the models in Tables 3–5.

To assess heterogeneity among the cohorts and the data entry programs, study-specific log odds ratios per quintile, as defined above, were weighted by the inverse of their variance, combined using fixed-effects meta-analysis, and evaluated by use of the I² statistic.

Absolute risks were calculated for EPIC-Norfolk (62% of the study participants). Individuals were excluded from the cohort if they had a registry-reported cancer at baseline (except nonmelanoma skin cancer), leaving 24,211 individuals for analysis. Of these, 412 had a colorectal cancer diagnosis between completion of the diary and the end of follow-up (December 31, 2006). Not all case patients were selected for the case–control study because some had been diagnosed within 12 months of completing the food diary and others did not have complete food diaries. Individuals were censored at death (other causes) and at any other cancer diagnosis (except nonmelanoma skin cancer). The absolute risk of colorectal cancer within 5 years of completing the food diary was estimated using a Kaplan–Meier analysis. The absolute risk of colorectal cancer for individuals in the lowest and highest quintiles of fiber intake was calculated using the observed proportion of case patients and control subjects in each of the lowest and highest quintiles from the case–control study. Quintiles are based on distribution of fiber intake among control subjects in EPIC-Norfolk. This analysis assumes that case patients and control subjects in the case–control study are representative of case patients and control subjects in the full cohort.

Measures from food diaries are subject to within-person random error, which result in attenuated odds ratio estimates (53,54). When repeat measures of intake from food diaries are available, regression calibration can be used to estimate within-person variation and hence to correct for the effects of measurement error on the odds ratio estimates associated with intake (53,55). Therefore, we used second 7-day food diaries that had been completed and analyzed to date (September 2009) by 411 (130 case patients and 281 control subjects) of the 1,590 participants from EPIC-Norfolk. We let \( R_i \) and \( R_j \) denote the two measures of intake from the paired food diaries and \( T \) denote unobserved true intake. Under the classical measurement error assumption, the measures of intake are related to \( T \) by \( R_i = T + e_i \) \((j = 1, 2)\), where the errors \( e_i \) \((j = 1, 2)\) are uncorrelated with each other and with \( T \). By making this assumption, we fitted univariate (fiber) and multivariate (fiber, energy from fat and nonfat, folate, and alcohol) regression calibration models (53,55) using the paired food diaries to give corrected odds ratio estimates. However, validation studies involving recovery biomarkers that were measured for some individuals in EPIC-Norfolk. For sodium, protein, and potassium intakes through recovery biomarkers that were measured for some individuals in EPIC-Norfolk. Sodium, protein, and potassium intakes, \( \beta \) was estimated to be 0.47, 0.81, and 0.69, respectively, and \( \rho \) was estimated to be 0.52, 0.52, and 0.58, respectively. We used the mean of these estimated \( \beta \) and \( \rho \) values across the three nutrients—0.66 and 0.54, respectively—to assess the potential impact of systematic and person-specific measurement error on odds ratio estimates in a univariate regression calibration. To perform a sensitivity analysis to assess the potential effects of systematic and person-specific errors in a multivariate regression calibration model, we would need fixed values of \( \beta \) and \( \rho \) for each of the five nutrients. For these analyses, we used the mean of the estimated \( \beta \) and \( \rho \) values among the five nutrients in the multivariate calibration (fiber, energy from fat, energy from nonfat sources, folate, and alcohol) and, additionally, parameters to account for person-specific errors that are correlated between nutrients as well as within nutrients (\( \rho \)). We did not attempt this here.

Two-sided \( P \) values less than .05 were considered statistically significant. All statistical analyses were done using Stata v.10 software (StataCorp, College Station, TX).

**Results**

A total of 579 incident colorectal cancer case patients and 1996 matched control subjects were available for analysis; 380 case patients had been diagnosed with colon cancer, and 199 had been diagnosed with rectal cancer. There were no statistically significant differences between case patients and control subjects with respect to mean values or distributions of participant weight; height; physical activity level; education level; smoking habits; socioeconomic status; or intakes of dietary fiber, energy, folate, or alcohol (data not shown). When stratified by quintiles of absolute dietary fiber intake, those in the highest quintile of fiber intake were younger and taller, had a lower body mass index, were more physically active, had attained higher levels of education, smoked less, and consumed more energy and folate but less alcohol compared with those in the lowest quintile of fiber intake (Table 2). When participants were stratified by quintiles of dietary fiber intake density, those with the highest quintile of fiber intake density per day were older, weighed less, had a lower body mass index, smoked less, and consumed more folate and less energy and alcohol compared with those with the lowest quintile of fiber intake density (data not shown). The main sources of dietary fiber for the 1590 participants from EPIC-Norfolk (62% of the study participants) were cereals (48.5%), followed by vegetables in the distribution of fiber intake among control subjects in EPIC-Norfolk. For sodium, protein, and potassium intakes, \( \beta \) was estimated to be 0.47, 0.81, and 0.69, respectively, and \( \rho \) was estimated to be 0.52, 0.52, and 0.58, respectively. We used the mean of these estimated \( \beta \) and \( \rho \) values across the three nutrients—0.66 and 0.54, respectively—to assess the potential impact of systematic and person-specific measurement error on odds ratio estimates in a univariate regression calibration. To perform a sensitivity analysis to assess the potential effects of systematic and person-specific errors in a multivariate regression calibration model, we would need fixed values of \( \beta \) and \( \rho \) for each of the five nutrients in the multivariate calibration (fiber, energy from fat, energy from nonfat sources, folate, and alcohol) and, additionally, parameters to account for person-specific errors that are correlated between nutrients as well as within nutrients (\( \rho \)). We did not attempt this here.

Two-sided \( P \) values less than .05 were considered statistically significant. All statistical analyses were done using Stata v.10 software (StataCorp, College Station, TX).

A total of 579 incident colorectal cancer case patients and 1996 matched control subjects were available for analysis; 380 case patients had been diagnosed with colon cancer, and 199 had been diagnosed with rectal cancer. There were no statistically significant differences between case patients and control subjects with respect to mean values or distributions of participant weight; height; physical activity level; education level; smoking habits; socioeconomic status; or intakes of dietary fiber, energy, folate, or alcohol (data not shown). When stratified by quintiles of absolute dietary fiber intake, those in the highest quintile of fiber intake were younger and taller, had a lower body mass index, were more physically active, had attained higher levels of education, smoked less, and consumed more energy and folate but less alcohol compared with those in the lowest quintile of fiber intake (Table 2). When participants were stratified by quintiles of dietary fiber intake density, those with the highest quintile of fiber intake density per day were older, weighed less, had a lower body mass index, smoked less, and consumed more folate and less energy and alcohol compared with those with the lowest quintile of fiber intake density (data not shown). The main sources of dietary fiber for the 1590 participants from EPIC-Norfolk (62% of the study participants) were cereals (48.5%), followed by vegetables including potatoes (27.1%); fruit (16.0%) and legumes, seeds, and nuts (8.4%).
Table 2. Baseline characteristics of case patients and control subjects by sex-specific quintile of fiber intake as assessed by food diaries*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Pt†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean intake, g/d (SD)</td>
<td>8.9 (1.6)</td>
<td>12.3 (0.9)</td>
<td>14.6 (0.9)</td>
<td>17.6 (1.3)</td>
<td>24.1 (5.4)</td>
<td></td>
</tr>
<tr>
<td>No. of case patients/No. of control subjects</td>
<td>13/299</td>
<td>12/299</td>
<td>91/400</td>
<td>115/599</td>
<td>118/599</td>
<td></td>
</tr>
<tr>
<td>Male, No. (%)</td>
<td>249 (46.7)</td>
<td>249 (47.9)</td>
<td>236 (48.1)</td>
<td>256 (49.8)</td>
<td>256 (49.5)</td>
<td>.9</td>
</tr>
<tr>
<td>Mean age, y (SD)</td>
<td>62.9 (9.4)</td>
<td>61.7 (9.3)</td>
<td>61.8 (8.8)</td>
<td>61.3 (9.7)</td>
<td>61.3 (10.2)</td>
<td>.03</td>
</tr>
<tr>
<td>Mean weight, kg (SD)</td>
<td>73.1 (13.3)</td>
<td>72.4 (12.4)</td>
<td>73.9 (13.3)</td>
<td>73.0 (13.5)</td>
<td>71.8 (14.7)</td>
<td>14</td>
</tr>
<tr>
<td>Mean height, m (SD)</td>
<td>1.66 (0.09)</td>
<td>1.66 (0.09)</td>
<td>1.68 (0.09)</td>
<td>1.68 (0.09)</td>
<td>1.69 (0.10)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mean body mass index, kg/m² (SD)</td>
<td>26.6 (4.1)</td>
<td>26.1 (3.6)</td>
<td>26.2 (3.7)</td>
<td>25.8 (3.8)</td>
<td>25.0 (4.0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Physical activity‡, No. (%)</td>
<td>204 (41.5)</td>
<td>174 (36.3)</td>
<td>151 (32.4)</td>
<td>130 (27.0)</td>
<td>130 (27.0)</td>
<td></td>
</tr>
<tr>
<td>Inactive</td>
<td>150 (30.5)</td>
<td>137 (28.5)</td>
<td>152 (32.6)</td>
<td>154 (32.0)</td>
<td>165 (34.3)</td>
<td></td>
</tr>
<tr>
<td>Moderately inactive</td>
<td>85 (17.3)</td>
<td>103 (21.5)</td>
<td>94 (20.2)</td>
<td>105 (21.8)</td>
<td>106 (22.0)</td>
<td></td>
</tr>
<tr>
<td>Moderately active</td>
<td>53 (10.8)</td>
<td>66 (13.8)</td>
<td>69 (14.8)</td>
<td>92 (19.1)</td>
<td>80 (16.6)</td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>262 (51.4)</td>
<td>205 (42.6)</td>
<td>159 (34.8)</td>
<td>172 (36.1)</td>
<td>151 (33.6)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Education level, No. (%)</td>
<td>262 (51.4)</td>
<td>205 (42.6)</td>
<td>159 (34.8)</td>
<td>172 (36.1)</td>
<td>151 (33.6)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No formal qualifications</td>
<td>75 (14.7)</td>
<td>63 (13.1)</td>
<td>79 (17.3)</td>
<td>70 (14.7)</td>
<td>60 (13.4)</td>
<td></td>
</tr>
<tr>
<td>Lower secondary school</td>
<td>127 (24.9)</td>
<td>140 (29.1)</td>
<td>140 (30.6)</td>
<td>135 (28.3)</td>
<td>132 (29.4)</td>
<td></td>
</tr>
<tr>
<td>Higher secondary school</td>
<td>46 (9.0)</td>
<td>73 (15.2)</td>
<td>79 (17.3)</td>
<td>109 (21.0)</td>
<td>106 (23.6)</td>
<td></td>
</tr>
<tr>
<td>University degree</td>
<td>96 (18.3)</td>
<td>55 (10.7)</td>
<td>41 (8.5)</td>
<td>24 (4.7)</td>
<td>23 (4.5)</td>
<td></td>
</tr>
<tr>
<td>Smoking status, No. (%)</td>
<td>224 (42.6)</td>
<td>235 (45.8)</td>
<td>202 (41.7)</td>
<td>233 (45.7)</td>
<td>219 (42.5)</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>206 (39.2)</td>
<td>241 (49.8)</td>
<td>241 (49.8)</td>
<td>253 (49.6)</td>
<td>273 (53.0)</td>
<td></td>
</tr>
<tr>
<td>Former</td>
<td>46 (13.1)</td>
<td>41 (11.4)</td>
<td>46 (12.7)</td>
<td>43 (11.5)</td>
<td>41 (9.5)</td>
<td>.6</td>
</tr>
<tr>
<td>Never</td>
<td>39 (7.9)</td>
<td>42 (8.5)</td>
<td>50 (10.5)</td>
<td>58 (11.9)</td>
<td>60 (12.2)</td>
<td></td>
</tr>
<tr>
<td>Uses aspirin, No. (%)</td>
<td>185 (37.2)</td>
<td>192 (39.0)</td>
<td>215 (45.2)</td>
<td>195 (40.1)</td>
<td>206 (42.0)</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>103 (20.7)</td>
<td>87 (17.7)</td>
<td>89 (18.7)</td>
<td>95 (19.6)</td>
<td>105 (21.4)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>95 (17.0)</td>
<td>102 (20.7)</td>
<td>62 (13.0)</td>
<td>75 (15.4)</td>
<td>68 (13.9)</td>
<td></td>
</tr>
<tr>
<td>III-NM</td>
<td>67 (13.5)</td>
<td>47 (9.5)</td>
<td>48 (10.1)</td>
<td>48 (9.9)</td>
<td>40 (8.2)</td>
<td></td>
</tr>
<tr>
<td>III-M</td>
<td>18 (3.6)</td>
<td>23 (4.7)</td>
<td>12 (2.5)</td>
<td>15 (3.1)</td>
<td>11 (2.2)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>2.52 (0.84)</td>
<td>2.84 (0.94)</td>
<td>2.91 (0.92)</td>
<td>2.99 (0.99)</td>
<td>3.01 (1.12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mean energy from fat, MJ/d (SD)</td>
<td>5.17 (1.27)</td>
<td>5.19 (1.30)</td>
<td>5.42 (1.20)</td>
<td>5.69 (1.29)</td>
<td>6.09 (1.40)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mean folate intake, µg/d (SD)</td>
<td>401 (76.0)</td>
<td>244 (42.6)</td>
<td>264 (58.2)</td>
<td>291 (63.0)</td>
<td>345 (66.0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mean alcohol intake, g/d (SD)</td>
<td>14.7 (21.3)</td>
<td>13.0 (18.2)</td>
<td>12.1 (16.9)</td>
<td>11.4 (16.3)</td>
<td>9.8 (14.6)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mean fiber intake density, g/MJ (SD)</td>
<td>1.3 (0.4)</td>
<td>1.6 (0.4)</td>
<td>1.8 (0.4)</td>
<td>2.1 (0.5)</td>
<td>2.8 (0.8)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

* — not applicable.
† Two-sided x² test of differences in covariates across quintiles of fiber intake.
‡ Inactive = sedentary job and no recreational activity; moderately inactive = sedentary job with less than 0.5 hour recreational activity per day or standing job with no recreational activity; moderately active = sedentary job with 0.5–1 hour recreational activity per day or standing job with less than 0.5 hour recreational activity per day or physical job with no recreational activity; active = sedentary job with more than 1 hour recreational activity per day or standing job with more than 1 hour recreational activity per day or physical job with at least some recreational activity or heavy manual job.
§ I = professional occupations; II = managerial and technical occupations; III-NM = skilled occupations, nonmanual; III-M = skilled occupations, manual; IV = partly skilled occupations; V = unskilled occupations.

In unadjusted analyses, the odds of developing colorectal cancer was 25% lower for those in the highest compared with the lowest sex-specific quintile of dietary fiber intake (OR = 0.75, 95% CI = 0.55 to 1.02) (Table 3). This association remained consistent in analyses that adjusted for participant height, weight, energy intakes from fat and nonfat sources, and alcohol and folate intakes (OR = 0.73, 95% CI = 0.49 to 1.09) and was independent of smoking habits, socioeconomic status, education level, and physical activity (OR = 0.67, 95% CI = 0.42 to 1.05) (data not shown). Further adjustment for dietary intakes of vitamin D, calcium, and red and processed meats did not alter the association (OR = 0.66, 95% CI = 0.42 to 1.05) (data not shown).

The study-specific odds ratios of colorectal cancer per quintile of fiber intake generally supported an inverse association, although most of the 95% confidence intervals included 1.0 (Figure 1, A). The summary odds ratio per quintile of fiber intake determined by fixed-effects meta-analysis was 0.91 (95% CI = 0.83 to 1.00), and there was no evidence of heterogeneity among centers or between data entry programs (I² = 0%, P = .907) (Figure 1, A).

Unadjusted analyses of the linear association between the risk of colorectal cancer and a 6-g/d increase in dietary fiber intake assessed by food diaries [approximately 1 SD of dietary fiber intake across the whole study and one-third of the recommended average daily intake for fiber NSP (58)] showed a similarly strong association...
Table 3. Odds ratios (and 95% confidence intervals) of colorectal cancer across sex-specific quintiles of daily fiber intake and fiber intake density ascertained from food diaries

<table>
<thead>
<tr>
<th>Quintile of intake</th>
<th>Exposure variable and cancer site</th>
<th>Modela,b,c</th>
<th>Odds ratio (95% CI)</th>
<th>P (P_{\text{trend}})</th>
<th>Odds ratio (95% CI)</th>
<th>P (P_{\text{trend}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mean fiber intake, g/d (SD)</td>
<td>a</td>
<td>0.68 (0.49 to 0.95)</td>
<td>0.052</td>
<td>0.87 (0.65 to 1.16)</td>
<td>0.086</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>b</td>
<td>1.00 (referent)</td>
<td></td>
<td>1.80 (0.2)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>c†</td>
<td>0.76 (0.52 to 1.11)</td>
<td></td>
<td>0.67 (0.42 to 1.05)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>b</td>
<td>0.76 (0.52 to 1.11)</td>
<td></td>
<td>0.67 (0.42 to 1.05)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>c†</td>
<td>0.76 (0.52 to 1.11)</td>
<td></td>
<td>0.67 (0.42 to 1.05)</td>
<td></td>
</tr>
<tr>
<td>Per quintile, (P_{\text{trend}})</td>
<td></td>
<td>a</td>
<td>0.68 (0.49 to 0.95)</td>
<td>0.052</td>
<td>0.87 (0.65 to 1.16)</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b</td>
<td>1.00 (referent)</td>
<td></td>
<td>1.80 (0.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c†</td>
<td>0.76 (0.52 to 1.11)</td>
<td></td>
<td>0.67 (0.42 to 1.05)</td>
<td></td>
</tr>
</tbody>
</table>

*Model adjusted for age, sex, smoking status, and physical activity.
† 443 Case patients and 1673 control subjects.
‡ 372 Case patients and 1298 control subjects.
§ 192 Case patients and 672 control subjects.

In analyses that further adjusted for anthropomorphic factors; smoking; education level; socioeconomic status; physical activity; and intakes of energy, alcohol, and folate, the association did not change substantially (OR = 0.84, 95% CI = 0.71 to 1.00, \(P_{\text{trend}} = 0.056\) (Table 4), and further adjustment for dietary intakes of vitamin D, calcium, and red and processed meats did not alter the association (OR = 0.84, 95% CI = 0.70 to 1.00, \(P_{\text{trend}} = 0.053\) (data not shown).

The results for incident colon cancer were similar to those for colorectal cancer. When adjusted for height, weight, alcohol intake, energy intakes from fat and nonfat sources, the odds ratio of colon cancer for a one-quintile increase in fiber intake showed evidence of an inverse association (OR = 0.90, 95% CI = 0.81 to 0.99, \(P_{\text{trend}} = 0.029\) (data not shown). However, further adjustment for folate intake weakened the association (OR = 0.91, 95% CI = 0.82 to 1.01, \(P_{\text{trend}} = 0.086\) (Table 3). The association between a one-quintile increase in fiber intake and risk of rectal cancer was non-statistically significant in either unadjusted (\(P_{\text{trend}} = .5\); data not shown) or multivariable-adjusted (\(P_{\text{trend}} = .7\); Table 3) analyses. The differences between the odds ratios of fiber intake with colon cancer and with rectal cancer were not statistically significant (\(P = .5\) for absolute fiber and \(P = .2\) for fiber intake density).

To investigate whether the ratio of fiber intake to overall energy intake (rather than the absolute energy-adjusted intake) was associated with risk of colorectal cancer, we determined the odds ratio of colorectal cancer by quintiles of sex-specific fiber intake density (grams per megajoule). The unadjusted odds ratio for the highest quintile of fiber intake density compared with that for the lowest quintile showed evidence of an association with reduced risk of incident colorectal cancer (OR = 0.80, 95% CI = 0.59 to 1.08) (Table 3). Adjustment for height, weight, energy intakes from fat and from nonfat sources, dietary intake of folate, and alcohol intake strengthened the association (OR = 0.66, 95% CI = 0.45 to 0.96) (Table 3). Further adjustment for participant smoking habits; physical activity; socioeconomic status; education level; and dietary intakes of vitamin D, calcium, and red and processed meats did not substantially change the association (OR = 0.62, 95% CI = 0.40 to 0.96) (data not shown). We also observed a strong association when we used a fully adjusted model to estimate the odds of colorectal cancer for a 0.7-g/MJ continuous increase in fiber intake (OR = 0.83, 95% CI = 0.70 to 0.97, \(P_{\text{trend}} = .018\) (Table 4), and replacing partitioned energy with total energy intake did not alter the association (OR = 0.85, 95% CI = 0.72 to 0.99, \(P_{\text{trend}} = .035\) (data not shown). The study-specific odds ratios per quintile of fiber intake density were similar to study-specific odds ratios per quintile of absolute fiber intake. The summary odds ratio per quintile of fiber intake density determined by fixed-effects meta-analysis was 0.88 (95% CI = 0.80 to 0.96), and there was no evidence of heterogeneity among centers (\(I^2 = 0\%\); Figure 1, B). In a multivariable analysis, fiber intake density was associated with the risk of colon cancer (OR for highest vs lowest quintile of intake = 0.60, 95% CI = 0.38 to 0.95) but not with the risk of rectal cancer (\(P_{\text{trend}} = .4\) (Table 3).

We conducted sensitivity analyses that excluded case patients who were diagnosed with colorectal cancer within 2 years of diary completion, and our results were not substantially altered, although
Figure 1. Forest plots of the within-center and pooled odds ratios (ORs) and 95% confidence intervals (CIs) for associations per quintile of fiber intake (grams per day) (A) and fiber intake density (grams per mega-joule) (B) with the risk of colorectal cancer. Odds ratios were adjusted for height, weight, energy from fat and nonfat sources, and dietary folate and alcohol intakes. The summary estimate was derived by fixed-effects inverse variance meta-analysis. The National Survey of Health and Development cohort was too small to include in the multivariable meta-analysis. Squares = study-specific odds ratios; size of the square = the weight given to this study (inverse of the variance of the log odds ratio) when estimating the summary odds ratio; horizontal lines = study-specific confidence intervals; diamond = summary estimate combining the study-specific estimates with a fixed-effects model; solid vertical line = odds ratio of 1; dashed vertical line = summary odds ratio. EPIC = European Prospective Investigation into Cancer and Nutrition; DANTE = Diet and Nutrition Tool for Evaluation; DINER = Data Into Nutrients for Epidemiological Research; UKWCS = UK Women’s Cohort Study.

When we used the residual method for energy adjustment in which we used the residuals of fiber intake from a linear regression of fiber measurements on energy measurements, the estimates of the association between residuals of fiber intake and the risk of colorectal cancer were similar to the ones we observed for absolute fiber intake and fiber intake density, adjusted for partitioned energy intake. For example, the odds ratio per quintile of fiber intake residuals, adjusted for height, weight, fat and nonfat energy, smoking, education level, socioeconomic status, physical activity,
### Table 4. Odds ratios (ORs) and 95% confidence intervals (CIs) of colorectal cancer associated with 1 SD increases in daily fiber intake and fiber intake density as ascertained from food diaries (564 case patients and 1970 control subjects) and from food-frequency questionnaires (FFQs) (483 case patients and 1784 control subjects)*

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Model</th>
<th>Food diary</th>
<th>FFQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>P</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Fiber intake, g/d</td>
<td>a</td>
<td>0.89 (0.79 to 0.99)</td>
<td>.034</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.86 (0.74 to 1.00)</td>
<td>.046</td>
</tr>
<tr>
<td></td>
<td>c†</td>
<td>0.84 (0.71 to 1.00)</td>
<td>.056</td>
</tr>
<tr>
<td>Fiber intake density, g/MJ</td>
<td>a</td>
<td>0.91 (0.82 to 1.00)</td>
<td>.059</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.83 (0.73 to 0.95)</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td>c†</td>
<td>0.83 (0.70 to 0.97)</td>
<td>.018</td>
</tr>
</tbody>
</table>

* A 1 SD increase corresponds to a 6-g/d increase in fiber intake and 0.7-g/MJ increase in fiber intake density.
† a = unadjusted; b = adjusted for height, weight, energy from fat and nonfat sources, and alcohol and dietary folate intakes; c = model b adjustments plus smoking status, education level, socioeconomic status, and physical activity.
‡ Food diary: 443 case patients and 1673 control subjects; FFQ: 420 case patients and 1615 control subjects.

and dietary intakes of alcohol and folate (comparable to Table 3, model c) was 0.89 (95% CI = 0.81 to 0.99, P<0.027) (data not shown).

Dietary data from FFQs that were completed before the food diary by participants (496 case patients and 1809 control subjects) in EPIC-Norfolk, EPIC-Oxford, the UKWCS, and Whitehall II were available for comparison. The mean fiber NSP intakes (SD) in case patients and control subjects as assessed by the FFQ were 19.7 (7.9) and 19.4 (7.3) g/d, respectively, which were higher than intakes assessed by food diaries in the same people (15.3 [5.8] and 15.6 [5.9] g/d, respectively). We observed no statistically significant association between absolute fiber intake in sex-specific quintiles as assessed by the FFQ and the risk of colorectal cancer in the unadjusted (P<0.2) or multivariable (P<0.1; Table 5) analysis, nor were statistically significant associations observed between fiber intake density in quintiles and the risk of colorectal cancer (unadjusted P<0.6, multivariable P<0.3; Table 5, Figure 2).

For example, comparing the highest with the lowest quintile of fiber intake density, the odds ratio was 0.88 (95% CI = 0.57 to 1.36) (model b in Table 5). Further adjustment for dietary intake of vitamin D, calcium, and red and processed meats did not change these results (absolute fiber intake: P<0.2; fiber intake density: P<0.2) (data not shown).

To assess the impact of measurement error on our results, we fitted a univariate regression calibration model that assumed classical measurement error by using data from 411 repeat 7-day diaries that were completed by 130 case patients and 281 control subjects (53). When we applied the correction that was used in the adjusted analysis (model c in Table 4), the corrected odds ratio for a 6-g/d increase in fiber intake was 0.77 (95% CI = 0.62 to 0.97) and the correct odds ratio for a 0.7-g/MJ increase in fiber intake density was 0.75 (95% CI = 0.61 to 0.91). A further univariate regression calibration that allowed for systematic and person-specific errors in measures of fiber intake resulted in a corrected odds ratio for a 6-g/d increase in fiber intake of 0.67 (95% CI = 0.31 to 0.95) and a corrected odds ratio for a 0.7-g/MJ increase in fiber intake density of 0.67 (95% CI = 0.31 to 0.95).
of dietary fiber decreases colonic transit time without altering the fermentation of fiber by the colonic microflora resulting in production of short-chain fatty acids that stimulate apoptosis (59). The association between dietary fiber intake and the risk of colorectal cancer was of a similar magnitude as that for the risk of colorectal cancer, but there was little evidence of an association between dietary fiber intake and the risk of rectal cancer. Similar results have previously been reported (15,51) and may reflect the fact that dietary fiber decreases colonic transit time without altering the storage time of stool in the rectum.

The relationship between dietary fiber intake and risk of colorectal cancer has been debated for many years (15–23). The recent WCRF/AICR report (24) found a dose–response effect of dietary fiber intake on colorectal cancer risk in a meta-analysis of existing cohort studies (RR per 10 g/d = 0.90, 95% CI = 0.84 to 0.97). However, residual confounding could not be ruled out, and the evidence was deemed probable rather than convincing. All of the studies included in the meta-analysis relied on data that were collected by FFQs, which have been used in large epidemiological studies to assess usual dietary intake of listed foods because they are easy to administer and have high return rates from participants. However, dietary intake data collected via FFQs may be crude because FFQs are restricted to a short list of some 100–200 items compared with the many thousands of foods in population food supplies. FFQs are retrospective, and there may be errors in assessing the frequency of consumption of foods and differences in perceptions of portion sizes. In addition, studies that have used biomarkers as objective measures of intake, such as doubly labeled water or urinary nitrogen, have shown that energy and protein intakes are poorly measured by FFQs (25,60–62), although measurement of protein intake density by FFQs appears to be considerably better (25,62). It is likely that intakes of other dietary components, including dietary fiber (28), are also poorly captured by FFQs, which would lead to an underestimate of their associations with cancer risk. In our analysis of food diary data, the estimated reduction in relative risk per 6 g/d fiber was greater than the apparent reduction in risk from 10 g fiber found in the WCRF/AICR report (Table 4) (24). Analytical methods for the determination of dietary fiber, the definitions of dietary fiber constituents, and, therefore, the quantitative recommendations for daily dietary fiber intake differ between countries, making comparisons between studies difficult (58,63). However, when we repeated these analyses in the pooled cohort using dietary data obtained from FFQs completed by the same individuals who had completed the food diaries and the same method for fiber analysis, we did not observe an association between fiber intake assessed by FFQ and the risk of colorectal cancer (Figure 2), despite the apparent higher intake of fiber NSP when intake was assessed by FFQ (Tables 3 and 5).

Although food diaries are probably better dietary assessment tools than FFQs, they do not completely eliminate measurement error. Therefore, by using repeat food diary measurements for 411 individuals from the EPIC-Norfolk study, we used regression calibration to correct for the effects of error in the food diary measurements of fiber intake. A multivariate regression calibration model that assumed classical measurement error resulted in corrected odds ratio estimates of 0.72 (95% CI = 0.51 to 1.02) for a 6-g/d increase in fiber intake (uncorrected OR = 0.84, 95% CI = 0.71 to 1.00), and 0.68 (95% CI = 0.48 to 0.96) for a 0.7-g/MJ increase in fiber density (uncorrected OR = 0.83, 95% CI = 0.70 to 0.97). Sensitivity
analyses that allowed for systematic and person-specific errors in univariate regression calibration suggested that the inverse association between fiber intake and the risk of colorectal cancer could be even stronger; however, without the use of unbiased measures of intake such as recovery biomarkers, it is not possible to assess this possibility further. Our sensitivity analysis included parameter values from a study that used recovery biomarkers for sodium, protein, and potassium, and it is possible that the values of parameters β and ρ may not be transferable across nutrients.

This study has four limitations. First, although our food diary data were mainly entered on the same processing program, DIER, some food diaries from UKWCS and all food diaries from Medical Research Council NSHD had previously been entered into other systems, which introduced an additional potential source of measurement error that could contribute to bias in the risk estimates. Second, not all of the participating studies employed 7-day food diaries, which may also introduce measurement error in the exposure. However, 4-day diaries showed good agreement with longer diaries for averaged nutrient intakes in our data. Third, our results refer to fiber in foods because we were unable to assess the use of fiber supplements.

Fourth, some of the participating cohorts recorded self-reported anthropometric data (36,38), whereas in other cohorts, these data were recorded by trained interviewers. Nevertheless, by matching case patients and control subjects within cohorts and using a matched analysis, we have reduced the potential effects of any differential errors resulting from different data collection procedures.

This study has several strengths. By nesting our case–control study within established prospective cohort studies, we have avoided the problems of recall bias and selection bias. Whereas logistic considerations preclude the analysis of all food diaries on a cohort level, case–control studies that are nested within cohorts and that match four control subjects to each case patient provide highly efficient odds ratio estimates compared with odds ratio estimates from whole cohort studies that have complete exposure information. Diary data entry in this study was standardized, and fiber values were derived from the same analytical method and set of food composition tables (34). The quintiles of intake were derived from the range of values for the entire pooled dataset, and each cohort contributed subjects to every quintile of intake. Sensitivity analyses that excluded participants who were diagnosed with colorectal cancer within the first 2 years of follow-up did not alter our results nor did adjustment for age. Finally, regression calibration for measurement error supported our primary results.

In summary, by using pooled data from mature prospective cohorts in the United Kingdom, we have shown a strong inverse association between dietary fiber intake assessed by a detailed food diary and the subsequent development of colorectal cancer. For individuals who consumed an average of 24 g per day of fiber NSP (the highest quintile), the odds of developing colorectal cancer were 30% lower than that for individuals who consumed an average of 10 g per day of fiber NSP (the lowest quintile); for individuals who consumed an average of 18 g per day of fiber NSP as recommended by the UK Department of Health (58), the odds of developing colorectal cancer were approximately 20% lower. Adjustment for dietary folate intake and other potential confounders did not alter these findings. The associations with colorectal cancer risk were stronger for fiber intake density than for absolute dietary fiber intake. These findings strengthen existing evidence that supports recommendations to increase dietary fiber intake in populations to reduce colorectal cancer incidence. The fact that we found no association using exposures assessed using a simpler method of dietary assessment, the FFQ, may explain the lack of convincing evidence relating fiber intake to a substantial reduction in colorectal cancer risk in some previous studies that relied on FFQs.

**References**


**Funding**

Medical Research Council (G.M., D.K., E.J.B.; g402517 and g0500300 to K.T.K., U.1052.00.006 to I.W.); research program grant funding from Cancer Research UK (C570/A5028 to T.J.K., C/864/A8257 to K.T.K.); the Lloyds TSB Charitable Foundation for the Channel Islands (I.S.F.); the World Cancer Research Fund (J.E.C., V.J.B.).

**Notes**

The study sponsors did not have any role in the design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

**Affiliations of authors:** Medical Research Council (MRC) Centre for Nutritional Epidemiology in Cancer Prevention and Survival, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK (CCD, RHK, SAR); MRC Biostatistics Unit, Institute of Public Health (RHK, IRW), University of Cambridge, Cambridge, UK (RHK, IRW); Cancer Epidemiology Unit, University of Oxford, Oxford, UK (EAS, TJK); Centre for Epidemiology and Biostatistics, University of Leeds, Leeds, UK (DCG, JEC, VJB); Academic Oncology Unit, Guy’s Hospital, London, UK (ISF); Department of Epidemiology and Public Health (MJS, EJB) and MRC Unit for Lifelong Health and Ageing, Department of Epidemiology and Public Health (GM, DK), University College London, London, UK; MRC Human Nutrition Research, Elsie Widdowson Laboratory, Cambridge, UK (AMS); Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK (RL, MAHL, KTK).