Cancer Epidemiol Biomarkers Prev; 19(10); 2534–40. ©2010 AACR.

Cruciferous Vegetable Intake and Lung Cancer Risk: A Nested Case-Control Study Matched on Cigarette Smoking

Tram Kim Lam1,5,6, Ingo Ruczinski2, Kathy J. Helzlsouer1,4,8, Yin Yao Shugart7, Laura E. Caulfield3, and Anthony J. Alberg1,9

Abstract

Background: Due predominantly to cigarette smoking, lung cancer is the leading cancer-related cause of death worldwide. Cruciferous vegetables may reduce lung cancer risk. The association between intake of cruciferous vegetables and lung cancer risk was investigated in the CLUE II study, a community-based cohort established in 1989.

Methods: We matched 274 incident cases of lung cancer diagnosed from 1990 to 2005 to 1,089 cancer-free controls on age, sex, and cigarette smoking. Dietary information was collected at baseline. Multivariable odds ratios (OR) and 95% confidence intervals (95% CI) were calculated using conditional logistic regression.

Results: Intake of cruciferous vegetables was inversely associated with lung cancer risk (highest-versus-lowest fourth: ORQ4vsQ1, 0.57; 95% CI, 0.38-0.85; P-trend = 0.01). The inverse associations held true for former smokers (ORQ4vsQ1, 0.49; 95% CI, 0.27-0.92; P-trend = 0.05) and current smokers (ORQ4vsQ1, 0.52; 95% CI, 0.29-0.95; P-trend = 0.02).

Conclusions: After carefully controlling for cigarette smoking, higher intake of cruciferous vegetable was associated with lower risk of lung cancer.

Impact: The observed inverse association coupled with accumulating evidence suggests that intake of cruciferous vegetables is inversely associated with lung cancer risk, and this association seems to hold true beyond the confounding effects of cigarette smoking.

Introduction

Cruciferous vegetables (e.g., broccoli, cabbage, mustard greens, Brussels sprouts) have generated interest as dietary constituents that may protect against lung cancer (1). These vegetables are a major source of glucosinolates, which are precursors for isothiocyanates and indole-3-carbinol (1, 2). These crucifer derivatives exhibit several anticarcinogenic properties. Consumption of broccoli decreased markers of DNA damage for both smokers and nonsmokers (3). Likewise, evidence showed that isothiocyanates and indole-3-carbinol inhibit the bioactivation of procarcinogens found in tobacco smoke (e.g., polycyclic aromatic hydrocarbons) into genotoxic carcinogens (4). Isothiocyanates, in particular, also induce the expression of glutathione S-transferase (GST) enzymes, which increase the excretion of carcinogens in urine (2, 5). Experimental evidence suggests that sulforaphane, a major isothiocyanate found in broccoli, can induce cell cycle arrest and apoptosis (6-8). By dually preventing activation and promoting inactivation of carcinogens and by exerting cell cycle control, isothiocyanates could protect against cancer. Furthermore, data have shown that sulforaphane influences epigenetic changes with consequences on cancer outcomes (9, 10).

We previously published a meta-analysis on cruciferous vegetable intake and lung cancer risk (11) and observed inverse associations in case-control studies [pooled odds ratio (OR), 0.78; 95% confidence interval (95% CI), 0.70-0.88] and cohort studies [pooled relative risk (RR), 0.83; 95% CI, 0.62-1.08]. The pooled results for cohort studies were based on six studies (12-16). Since then, the NIH-AARP Diet and Health Study, a large prospective cohort study, observed a RR of 0.92 (95% CI, 0.83-1.02; P-trend = 0.09) in men in the highest versus the lowest category of cruciferous vegetable intake, but none in women (RR, 1.02; 95% CI, 0.83-1.16; ref. 17). Additionally,
Cruciferous Vegetable Intake and Lung Cancer Risk

Potential confounding effects of cigarette smoking in these previous studies interfere with drawing clear-cut inferences. Cigarette smoking is the predominant cause of lung cancer (19), and cigarette smokers have significantly lower vegetable consumption than never smokers (20). Under these circumstances, the possibility remains that the inverse associations observed may have been due to residual confounding by smoking (21). Whereas all previously published cohort studies reported estimates that were at least minimally adjusted for smoking status, a lingering issue is the degree to which cruciferous vegetables may be associated with lung cancer risk beyond the potential confounding effects of cigarette smoking. The present study was designed to directly address this challenge. Specifically, the present study was carried out to investigate whether or not consumption of cruciferous vegetables was associated with reduced risk of lung cancer after tightly matching cases and controls to be similar with respect to cigarette smoking. We conducted a nested case-control study within a community-based, prospective cohort in Washington County, Maryland.

Materials and Methods

Study population

Established in 1989, the CLUE II cohort was named for its campaign slogan, “Give Us a Clue to Cancer and Heart Disease.” The details of the establishment of the cohort have been reported elsewhere (22, 23). Briefly, from May 1 through November 30, 1989, 32,897 residents of Washington County, Maryland, and surrounding area (30-mile radius) consented to participate in CLUE II. The population for the present study was restricted to 22,631 adults (>18 years old) with a Washington County address and therefore were covered by the Washington County Cancer Registry.

At baseline, participants completed a brief questionnaire on basic demographic characteristics, smoking status, and number of cigarettes smoked per day. At that time, participants were asked to complete and mail in a modified Block food frequency questionnaire (FFQ; ref. 24). Starting in 1996, the CLUE II cohort members received periodic follow-up questionnaires.

Ascertainment of lung cancer cases was achieved through linkage with the Washington County Cancer Registry, the Maryland State Cancer Registry, and the Maryland state mortality records. The Washington County Cancer Registry obtains its data primarily from the county’s only general hospital, Washington County Hospital, and its pathology department. Since 1992, CLUE II cohort members were also linked to the Maryland State Cancer Registry, which has a mandatory cancer reporting policy. Lastly, state mortality records of Washington County residents were consulted.

Case ascertainment was believed to be reasonably complete. Based on the most recent report from the National Program of Cancer Registries for 2006, the Maryland State Cancer Registry was considered 98% complete, and the Washington County Cancer Registry has been observed to have higher ascertainment than the state registry (25). Through 1998, the annual rate of loss to follow-up was estimated to be less than 1% based on follow-up of a random sample of cohort members (26). A current estimate was that 6.9% of the CLUE II participants may have moved out of the county. Thus, the study population was relatively stable.

Case and control selection

From May 1, 1990 through October 31, 2005, 274 incident lung cancer cases occurred within the CLUE II cohort. Cases were first-time diagnoses of lung cancer [International Classification of Diseases, Eighth Revision (ICD-8) and ICD-9), code 162 for cases that occurred from 1992 to 2000 and ICD-10 codes C33-C34 for cases diagnosed from 2001 to present). For each case, up to four controls were selected (one case had two controls and four cases had three controls). Eligibility criteria for controls were (a) completion of the baseline FFQ; (b) having no prior history of cancer except for nonmelanoma skin cancer or cervical cancer in situ; and (c) being cancer-free and alive at the time of case diagnosis.

Matching criteria and smoking exposure

Controls were individually matched to cases on the following variables: sex, race (white, black, other), age (±5 years), and smoking status (never, former, or current smokers). For former and current smokers, cases and controls were further matched on the number of cigarettes smoked per day (CPD) as follows. Ever smokers were categorized into three smoking groups: ≤19, 20 to 29, and ≥30 CPD. For those who smoked <30 CPD, cases and controls were matched within ±5 CPD. Those who smoked between 30 and 45 CPD were matched within ±10 CPD, and the heaviest smokers (≥45) were matched within ±20 CPD. The 1996 follow-up questionnaire was relevant to this study, as it provided additional details on pack-years of smoking in cases and controls. Among those (64%) who provided information on pack-years of smoking in the 1996 follow-up questionnaire, we further matched cases to controls within ±5 pack-years of cigarette smoking. Of the total number of lung cancer cases included in the study, 99 were diagnosed before 1996.

Dietary measurements

At baseline, a validated 60-item Block FFQ was mailed to all participants (24, 27). For each food, participants were asked to classify their usual intake during the previous year into one of nine categories of frequency of consumption, ranging from “never or less than once per month” to “2+ per day”. Respondents who ate a food were asked to categorize the average serving size as small, medium (defined as one-half cup), or large.

Based on these data, daily intake of each food item was calculated by multiplying the serving size by the frequency.
of consumption (expressed as intake per day). Measures of servings per day were calculated for intakes of (a) total vegetables, (b) total cruciferous vegetables, (c) total noncruciferous vegetables, and (d) total fruit.

Intakes of dietary food groups were all energy adjusted by the nutrient density method, in which daily dietary intake was expressed as servings per 1,000 kcal. In this study, total vegetables included broccoli, spinach, mustard greens/turnip greens/collards, cole slaw/cabbage/sauerkraut, green salad, carrots or mixed vegetables containing carrots, vegetable soups, and other vegetables (e.g., green beans, corn, peas). Total cruciferous vegetables intake included broccoli, cole slaw/cabbage/sauerkraut, and mustard greens/turnip greens/collards. Noncruciferous vegetables included all vegetables except broccoli, cole slaw/cabbage/sauerkraut, and mustard greens/turnip greens/collards. Total fruits included tomatoes (e.g., tomato juices and pastas such as spaghetti and lasagna with tomato sauce), apples/apple sauce/pears, cantaloupe, oranges, orange juice, grapefruit, and other fruits (e.g., bananas).

**Statistical analysis**

Quartiles of the distribution of the controls for each dietary group were used to classify study participants based on their intakes of (a) total vegetables, (b) cruciferous vegetables, (c) noncruciferous vegetables, and (d) fruits. At baseline, 7.5% (36 cases and 66 controls) of study participants had incomplete or no dietary information. Exclusion of these individuals from our study could introduce bias (28). To address the issue of incomplete dietary data in this study, we substituted missing data with the median value of the controls for each dietary factor.

Conditional logistic regression was used to estimate ORs and 95% CIs for the associations between total cruciferous vegetable intake or other dietary factors and lung cancer incidence. Unless otherwise indicated, all analyses were adjusted for body mass index (categories: <24.9, 25.0-29.9, and 30+ kg/m²), total energy intake (continuous), and total fruit and noncruciferous vegetable intakes (continuous). To minimize the possibility of residual confounding, we further adjusted for age and number of cigarettes smoked per day in the full models as continuous variables. Additional adjustment for total meat intake did not markedly alter the results and thus total meat intake was not included in these analyses.

Furthermore, stratified analyses were done by smoking status (never, former, and current smokers). Tests for dose-response trends across quartiles of dietary intake were estimated by fitting the ordinal exposure variables as ordered categories. All analyses were done using STATA version 9.1.0. A two-sided P value of <0.05 was considered to be statistically significant.

To investigate the possible effect of how the data for those with missing dietary information were handled, sensitivity analyses were conducted in two ways: (a) imputing missing values using the multiple imputation method (28, 29) and (b) excluding those with incomplete or missing data. For the imputation analysis, the estimates and variances from three imputed sets were combined to derive the reported total estimates and confidence intervals.

**Results**

After a mean follow-up of 15 years, 274 incident lung cancer cases (27% adenocarcinoma, 24% squamous cell lung cancer, 18% small cell lung cancer, and 31% others, including large cell lung cancer and nondifferentiated lung cancer) were diagnosed in CLUE II. For both cases and controls, compared with those in the highest fourth of cruciferous vegetable intake, those in the lowest fourth were slightly younger, less educated, more likely to smoke, and more likely to be current smokers (Table 1). Spearman correlation coefficients between cruciferous vegetable intake and noncruciferous vegetable, fruit, combined vegetable and fruit, and combined meat intakes were 0.58, 0.45, 0.62, and 0.21, respectively, indicating that cruciferous vegetable intake was associated with higher intake of fruits and other vegetables and lower meat intake.

The OR for the highest-versus-lowest fourth of cruciferous vegetable intake was 0.57 (95% CI, 0.38-0.85; P-trend = 0.01), after adjusting for age, number of cigarettes smoked per day, total energy intake, and total fruits and noncruciferous vegetables (Table 2). Inverse associations were observed for both men and women in the highest fourth of intake, but were statistically significant only for women (OR Q4 vs Q1 = 0.52; 95% CI, 0.29-0.92; P-trend = 0.07; versus for men: OR Q4 vs Q1 = 0.72; 95% CI, 0.37-1.37; P-trend = 0.18). Similar inverse associations were observed for all three major subtypes of lung cancer (adenocarcinoma, squamous cell lung cancer, and small cell lung cancer); however, none of the associations were statistically significant (data not shown).

After stratification by smoking status, inverse associations were observed in former smokers (OR Q4 vs Q1 = 0.49; 95% CI, 0.27-0.92; P-trend = 0.05) and current smokers (OR Q4 vs Q1 = 0.52; 95% CI, 0.29-0.95; P-trend = 0.02; Table 2). Conversely, the associations among never smokers were in the direction of increased risk; however, because of the small number of never smokers in this study, even when the second to fourth quartiles were combined, the association was very imprecise and compatible with chance (OR, 4.9; 95% CI, 0.63-37.8).

When cruciferous vegetable intake for individuals with missing dietary data was imputed using the multiple imputation method, those in the highest-versus-lowest quartile of cruciferous vegetable intake had a 23% lower risk of lung cancer (Supplementary Table S1). For the same comparison, the OR Q4 vs Q1 was 0.85 (0.52-1.39) when individuals with no dietary information were excluded.

**Discussion**

In this nested case-control study, we carefully controlled for cigarette smoking in the study design to investigate the association between cruciferous vegetable intake and lung cancer.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cases (n = 274)</th>
<th>Controls (n = 1,089)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q1 (lowest)</td>
<td>Q2</td>
</tr>
<tr>
<td>Total cruciferous vegetable intake (±SD)* †</td>
<td>0.08</td>
<td>0.18</td>
</tr>
<tr>
<td>Vegetables (±SD) † ‡</td>
<td>0.32 ± 0.57</td>
<td>1.06 ± 0.80</td>
</tr>
<tr>
<td>Fruits (±SD) §</td>
<td>0.36 ± 0.95</td>
<td>1.23 ± 1.14</td>
</tr>
<tr>
<td>Vegetables + fruits (±SD) † ‡</td>
<td>0.70 ± 1.36</td>
<td>2.40 ± 1.47</td>
</tr>
<tr>
<td>Mean age (y)</td>
<td>57.94 ± 11.75</td>
<td>60.05 ± 7.36</td>
</tr>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>35.56</td>
<td>40.00</td>
</tr>
<tr>
<td>Male</td>
<td>64.44</td>
<td>60.00</td>
</tr>
<tr>
<td>Mean education (grade)</td>
<td>11.19 ± 25.00</td>
<td>11.52 ± 2.43</td>
</tr>
<tr>
<td>Mean cholesterol</td>
<td>213.86 ± 43.44</td>
<td>209.45 ± 34.22</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>&lt;24.9 (%)</td>
<td>32.22</td>
</tr>
<tr>
<td>25.0-29.9 (%)</td>
<td>40.00</td>
<td>51.67</td>
</tr>
<tr>
<td>30+ (%)</td>
<td>27.78</td>
<td>11.67</td>
</tr>
<tr>
<td>Smoking status, case/control∥</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never, 20/80 (%)</td>
<td>5.56</td>
<td>6.67</td>
</tr>
<tr>
<td>Former, 110/440 (%)</td>
<td>35.56</td>
<td>35.00</td>
</tr>
<tr>
<td>Current, 144/569 (%)</td>
<td>58.89</td>
<td>58.33</td>
</tr>
</tbody>
</table>

NOTE: Intake = daily frequency of servings × portion size.
*Cruciferous vegetables: summary measure broccoli, cole slaw/cabbage/sauerkraut, and mustard greens/turnip greens/collards.
†Median number of servings per day per 1,000 kcal.
‡All vegetables: summary measure of tomatoes, broccoli, spinach, cole slaw/cabbage/sauerkraut, green salad, carrots, mustard greens/turnip greens/collards, vegetable soups, and other vegetables.
§Fruits: summary measure of tomato, spaghetti/lasagna/pasta containing tomato sauce, apples/applesauce/pears, oranges, cantaloupes, grapefruits, orange/grapefruit juices, and other fruits.
∥Number of cases and controls.
cancer risk. Overall, we observed statistically significant inverse associations between higher intake of cruciferous vegetable intake and lung cancer risk, independent of cigarette smoking. Similar significant inverse associations were observed for former and current smokers, but not for never smokers.

In the present study, we report a 43% lower risk of lung cancer comparing the highest versus the lowest category of cruciferous vegetable intake. We previously reported the results of a systematic review of the evidence on cruciferous vegetable intake and lung cancer from both cohort and case-control studies, which showed a pooled relative risk of 0.83 (95% CI, 0.62-1.08) for cohort studies (11). Inclusion of the results from the present study to previous cohort studies showed a pooled relative risk of 0.79 (highest-versus-lowest category: 95% CI, 0.61-1.01). Taken together, the epidemiologic evidence suggests that cruciferous vegetable intake may be inversely associated with lung cancer risk. The results of our study add to this evidence by indicating that the previously observed associations are unlikely to be due to the residual confounding effects of cigarette smoking.

Isothiocyanates and indole-3-carbinol are derivatives of glucosinolates found in cruciferous vegetables. These compounds individually and in combination may reduce the risk of lung cancer through multiple anticarcinogenic mechanisms (30). Most published reports had largely attributed the anticarcinogenic properties of cruciferous vegetables to isothiocyanates, although indole-3-carbinol has recently been shown to have chemopreventive characteristics (31). Previous studies (18, 32-35), including a systematic review (11), have suggested that GSTM1 and GSTT1 might modify the association between cruciferous vegetables and lung cancer risk. GSTM1 and GSTT1 are part of the GST family and are involved in isothiocyanate metabolism (2).

The present study has several methodologic strengths. Compared with other cohort studies, the hallmarks of the present study are that it was a nested case-control study carefully matched on cigarette smoking history, it had the longest duration of follow-up years (15 years), and it used a community-based cohort. The most unique aspect of the study for this topic is that cases and controls were well matched on several smoking characteristics (smoking status and number of cigarettes smoked per day) to minimize the strong confounding effect of cigarette smoking. The prospective nature of the dietary and smoking data minimizes the issue of recall bias by disease status.

There are, however, several limitations to this study. A small percentage of the study participants (<8%) had

Table 2. ORs and 95% CIs of total cruciferous vegetable intake and risk of lung cancer in Washington County, Maryland (1989-2005)

<table>
<thead>
<tr>
<th>Quartiles of cruciferous vegetables intake</th>
<th>P-trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td></td>
</tr>
</tbody>
</table>

| All | Case/control | 90/258 | 60/257 | 52/258 | 72/316 | 1.00 (reference) | 0.60 (0.41-0.89) | 0.53 (0.35-0.80) | 0.57 (0.38-0.85) | 0.01 |
| OR (95% CI)* | | | | | | | | | | |

| Sex | Case/control | 32/82 | 24/105 | 22/108 | 46/199 | 1.00 (reference) | 0.51 (0.27-0.97) | 0.47 (0.24-0.89) | 0.52 (0.29-0.92) | 0.07 |
| OR (95% CI)* | | | | | | | | | | |

| Male | Case/control | 58/176 | 36/152 | 30/150 | 26/117 | 1.00 (reference) | 0.67 (0.41-1.11) | 0.59 (0.34-1.03) | 0.72 (0.37-1.37) | 0.18 |
| OR (95% CI)* | | | | | | | | | | |

| Smoking status | Case/control | 5/27 | 4/14 | 6/16 | 5/23 | 1.00 (reference) | 4.04 (0.40-40.6) | 14.1 (0.89-221.74) | 4.52 (0.40-50.82) | 0.28 |
| OR (95% CI)* | | | | | | | | | | |

| Never smokers | Case/control | 32/73 | 21/99 | 22/116 | 35/152 | 1.00 (reference) | 0.44 (0.23-0.86) | 0.39 (0.20-0.76) | 0.49 (0.27-0.92) | 0.05 |
| OR (95% CI)* | | | | | | | | | | |

| Former smokers | Case/control | 53/158 | 35/144 | 24/126 | 32/141 | 1.00 (reference) | 0.61 (0.37-1.03) | 0.49 (0.27-0.88) | 0.52 (0.29-0.95) | 0.02 |
| OR (95% CI)* | | | | | | | | | | |


*Adjusted for age, body mass index, number of cigarettes smoked per day, energy intake (continuous), and total fruit and noncruciferous vegetable intakes (continuous).
missing data due to either incomplete questionnaire or not returning a dietary questionnaire. Exclusion of these subjects could introduce selection bias and might lead to erroneous inferences. We instead substituted the median values based on the controls distribution of each missing dietary factor. To assess the effect of this imputation approach, we performed sensitivity analyses using other approaches to address missing or incomplete data: (a) the multiple imputation method and (b) inclusion only of study participants with complete dietary data. For similar categorical comparisons, the results based on the sensitivity analyses showed weaker nonsignificant inverse associations. Thus, even with the relatively small percentage of missing data in the present study, the approach to handling this missing data can affect the inferences to a nontrivial degree, and the approach adopted in our primary analyses was less conservative than the other approaches. In the present study, we matched as closely as possible on smoking status and smoking history such that smoking is more strongly controlled than in previous studies on this topic. Nevertheless, the possibility of residual confounding by cigarette smoking cannot be completely eliminated. For example, the data on duration of smoking were incomplete for 36% of the cohort population, and thus we could only match cases and controls on pack-years in a subgroup of subjects. The distribution of histologic types of lung cancer in the present study differs slightly from the Surveillance Epidemiology and End Results data (36), with slightly higher proportions of squamous cell and small cell lung cancers and a lower proportion of adenocarcinoma. It is uncertain whether this is due to chance, a true population-based difference either in risk factors or diagnostic practices, or other factors, but this should not have had a major effect on the observed inverse associations between cruciferous vegetable consumption and lung cancer risk.

Lastly, although the Block FFQ had been validated and found to have adequately estimated dietary intake of Americans overall (27, 37), dietary data derived from FFQs are subject to measurement error that may be random or systematic (38). Dietary intake was obtained before lung cancer diagnosis in the present study; thus, the measurement error would most likely be nondifferential and would most likely have attenuated the associations toward the null. The individual vegetables (broccoli, cole slaw/cabbage/sauerkraut, and mustard greens/turnip greens/collards) that were included in our measure of cruciferous consumption did not cover the full breadth of cruciferous vegetables. Thus, we likely underestimated cruciferous vegetable intake in our population. However, in the United States, broccoli is the most commonly consumed cruciferous vegetable and, therefore, the major source of isothiocyanates (39). Particularly, in a population with low dietary cruciferous vegetable intake, such as the CLUE II cohort, the cruciferous vegetables included in our summary measure probably represented the crucifers commonly eaten in this community.

In summary, after carefully matching lung cancer cases and controls by smoking history, we observed statistically significant inverse associations between consumption of cruciferous vegetables and lung cancer risk that were closely aligned with the previous body of evidence on this topic. Taken together with the existing evidence from prospective cohort studies, at present, the totality of the evidence suggests that it is tenable that cruciferous vegetable intake is inversely associated with lung cancer risk, and that this association holds true beyond the confounding effects of cigarette smoking.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank the participants of the CLUE II cohort and the staff of the George W. Comstock Center for Public Health Research and Prevention for their dedication and contributions to the study. This study is dedicated to the memory of Dr. George W. Comstock.

Grant Support

Institutional Research Epidemiology Fellowship grant T32CA099314, National Institute of Aging grant 5U01AG08033, and National Cancer Institute grant CA105069. This report is based, at least in part, on information provided by the Maryland Cancer Registry, Maryland Department of Health and Mental Health.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 05/05/2010; revised 07/07/2010; accepted 07/23/2010; published OnlineFirst 09/14/2010.

References


