Carrots, Green Vegetables and Lung Cancer: A Case-Control Study

PAOLA PISANI*, FRANCO BERRINO*, MAURIZIO MACALUSO*, UGO PASTORINO†, PAOLO CROSIGNANI* AND ALBERTO BALDASSERONI*


A total of 417 lung cancer cases and 849 controls were interviewed on their life-long tobacco usage and their current intake of four food items rich in retinol or carotene. The study was a hospital-based case control where 'cases' were lung cancer patients diagnosed during the period 1979/80 at seven hospitals in the Lombardy region (90% pathologically confirmed) and controls were patients admitted to the same hospitals for causes unrelated to tobacco smoking (epithelial cancers being excluded from present analysis). Odds ratios (OR) have been computed for increasing frequencies of consumption of liver, cheese, carrots and leafy green vegetables, having controlled for the confounding effects of tobacco usage, residence and birthplace.

Current smokers who did not consume carrots showed a three-fold risk of developing lung cancer compared with those who ate them more than once a week (OR = 2.9 < p < 0.01); the ORs for consumers in the categories of 1 - 2 and 3 - 4 times per month were 1.8 and 2.0 respectively, with a significant test for linear trend (p < 0.01). Among ex-smokers or non-smokers, no decrease of lung cancer risk is evident associated with carrot consumption. An excess risk was also associated with low intake of green vegetables although it was not significant, while no excess risk was evident for non-consumers of liver and cheese. The effect of carrots is independent of histological type of lung cancer while the effect of green vegetables was confined to epidermoid carcinomas: low versus high intake group OR = 1.3.

A number of epidemiological studies have shown that a frequent intake of foods rich in vitamin A or its precursor beta-carotene is associated with a lower cancer risk in several different populations.1-12

The association between serum retinol and subsequent risk of cancer is not clear: four prospective studies have been reported. Two failed to observe lower levels of serum retinol among subjects who later developed cancer compared with a sample of subjects drawn from the same cohort.13,14 A third study reported a significant inverse association between serum retinol and subsequent risk of cancer.15 The strict negative association which resulted from the fourth investigation has been questioned: the analysis of data after a longer follow-up could not confirm the previous report.16,17 On the other hand there is strong evidence that a number of retinoids are effective in preventing the occurrence of cancer in various experimental systems.18

The association observed in humans is particularly consistent for lung cancer.3-12 Most of these studies considered tobacco smoking as a possible confounding factor. Generally, however, a rather rough categorization of tobacco consumption was used, therefore residual confounding cannot be excluded.

This paper presents data from a case-control study relating tobacco smoking to lung cancer. Within this framework information on a few food items was collected and the association with lung cancer risk could be evaluated with accurate data on smoking habits.

MATERIALS AND METHODS

In the context of an international project described elsewhere,19 lung cancer cases and hospital controls where interviewed about their life-time usage of tobacco. Detailed information was collected on amount, duration and brands of cigarettes, cigars and pipe tobacco consumed. The questionnaire included items on place of birth, residence, education, occupational history and alcohol consumption. Interviews were carried out from 1978 to 1981. During the last year of the study four new questions were introduced about intake of foods rich in vitamin A. They concerned monthly consumption of liver and carrots and weekly consumption of leafy green vegetables and cheese, expressed as usual frequencies of intake. The different...
centres participating in the project adopted different questions according to the typical diet of their own country. This analysis is restricted to cases and controls recruited after the relevant questions were introduced.

Cases and controls were enrolled in hospitals serving two different parts of Lombardy: three hospitals in Milan and four in the Varese Province. All cases were newly diagnosed lung cancers admitted to departments of pneumology or thoracic surgery in these hospitals. Three hundred and eighty-five out of 417 (92.3%) of the diagnoses were histologically or cytologically confirmed. Controls were selected from patients admitted to the same hospitals for causes unrelated to tobacco smoking. The original protocol required the interview of two controls per case matched by sex, age (± 5 years) and hospital. For the series recruited in Milan matching by area of residence was also required, while in the Varese province enrolment of cases and controls was restricted to residents. To avoid a long delay between the interview of cases and their matched controls, the latter were recruited independently of the occurrence of cases but according to the expected age and sex-specific pattern of lung cancer incidence. By this procedure, at the end of the study 1020 controls were available, 20% more than the requested number.

One hundred and seventy-one controls were excluded because they were affected by malignant neoplasms of epithelial origin. An unmatched analysis was thus performed on 417 cases (386 males) and 849 controls (785 males). Table 1 shows the distribution of diagnoses of controls together with the frequency of histological types of lung cancer.

A preliminary analysis according to percentiles of estimated retinol and beta-carotene indices of consumption has already been reported. Adjusted odds ratios (OR) and chi-square test for statistical significance were firstly computed by the Mantel–Haenszel procedure for data stratified according to potential confounders. Subsequently logistic modelling was used to allow simultaneous control of several factors. With the exception of age, continuous variables have been transformed into scaled variables so that only indicator variables entered the models. Maximum likelihood estimates of coefficients were obtained by means of the computer program PLR of the BMDP statistical package. Significance of ORs was assessed comparing the logarithm of OR divided by its estimated standard error with the unit normal distribution. Test for linear trend of the logarithms of ORs by categories of increasing consumption was performed testing significance of the coefficient of scaled variables defining the same categories of consumption considered in estimating odds ratios, but included in the model as interval variables. Approximate 95% confidence intervals of the logistic parameters were computed based on the normal distribution and on standard errors based on maximum likelihood estimation.

RESULTS
Table 2 summarizes a few characteristics of cases and controls by sex and by area of recruitment. As the proportion of controls excluded because they were affected by epithelial cancer was higher in Milan hospitals, the case-control ratio is different between the two groups; therefore area of recruitment was always considered in the analysis. The two study areas are similar for several relevant variables: proportion of tobacco smokers, age and education. On the contrary the two areas show different patterns for place of birth and residence. The lower proportion of cases born in southern Italy among subjects recruited in the Varese area is consistent with the lower incidence of lung
cancer observed by the local cancer registry among immigrants from the south.\textsuperscript{24} The lack of this difference in Milan is largely due to our Institute to which patients from all over the country are admitted. Since immigrants are known to have different dietary habits, geographical origin was controlled for in estimating ORs by logistic regression.

The overall association between frequency of consumption of the four food items under study and lung cancer risk is given in the last column of Table 3, showing ORs adjusted by study area and sex, with reference to the highest intake categories. They show weak effects of liver and cheese. On the contrary ORs for never consuming carrots or vegetables versus the corresponding high intake groups are, respectively, 2.0 and 1.9 and the trend by amount is appreciable.\textsuperscript{46}

Control for tobacco smoking was the next step in data analysis; OR estimates appeared quite heterogeneous among never smokers, ex-smokers and current smokers, particularly for carrot consumption, thus analysis was performed separately for each category.

Table 3 gives the distributions of cases and controls by smoking status and consumption of liver, cheese, carrots and leafy vegetables. Two estimates of the ORs are given for ex-smokers and current smokers: the first one is the Mantel–Haenszel estimate standardized by number of cigarettes per day, duration of smoking habit, study area and sex. The second estimate was obtained by fitting data with a logistic model including terms for liver, cheese, carrots, green vegetables, age, birthplace, residence, study area, number of cigarettes smoked per day, years as smoker and sex. Estimates for tobacco variables obtained from the same models are given below. Liver and cheese do not appear to be associated with the risk of lung cancer; ORs do not deviate significantly from unity in any of the three smoking categories, and nor is there an appreciable trend by increasing consumption. Estimates (standardized only by study area and sex) for non-smokers are very unstable due to the small number of cases, for example: logistic estimate of the OR of non-consumers of liver versus consumers is 1.2 (95\% c.l. (0.5, 2.8)), the estimate for non-consumers of carrots versus consumers is 0.9 (95\% c.l. (0.4, 2.2)).

Current smokers who do not consume carrots show a risk of developing lung cancer 2.9 times the risk for those who eat carrots more than once per week (95\% c.l. (1.6, 5.4)). Test for linear trend is significant (p<0.01). No excess risk is apparent among non-smokers or former smokers: OR associated with the lowest intake group versus the highest one among ex-smokers is 0.9 (95\% c.l. (0.3, 2.2)).

Controlling for education, age at which smoking started and use of filtered/unfiltered cigarettes was also attempted but none of these variables improved goodness of fit nor did they affect the estimated risks for the dietary items.

Excess risk associated with lower intake levels of vegetables remains after adjustment by confounders; some excess risk is still present among both current and ex-smokers. None of the estimates regarding vegetables is statistically significant. Among current smokers the logistic estimates are lower than those obtained by the Mantel–Haenszel procedure: this is due to standardization by carrot consumption.

ORs associated with carrot consumption have been analysed among former smokers by duration of not smoking: in none of the three categories of ex-smoking (1–5 years, 6–14 years, 15 or more years of not smoking).
smoking), did carrots show any protective effect.

Table 4 gives the logistic estimates of ORs for different levels of consumption of carrots and green vegetables by morphological type, adjusted for age, study area, places of birth and residence, smoking habit (smoking status, duration and daily amount) and sex. The protective effect of green vegetables is confined to epidermoid carcinomas, the OR for the lowest intake category is 1.3 (95% c.i. (0.7, 2.5)). Trend by increasing consumption does not reach statistical significance.
The effect is present among both smokers and the group of former smokers and non-smokers pooled.

Carrots on the contrary, show protection for all histological types, the trend is statistically significant for both epidermoid carcinomas and microc continues. The effect seems somewhat reduced and is not significant for adenocarcinomas. For reasons of readability data are not presented separately by smoking status. However, for each morphological type, the carrot effect was present only among current smokers as in Table 3.

**DISCUSSION**

The results presented showing a protective effect of carrots and green vegetables, are consistent with other case-control observations and with the three published cohort studies in which dietary exposure has been investigated. Lack of an effect of liver and cheese is also consistent with a few reports in which retinol sources were distinguished from dietary carotene. In our data, however, the pattern of protection is different according to tobacco smoking and histological type, ie carrots seem to be effective independent of histological type but only among current smokers, while green vegetables show a protective effect only for squamous cell carcinomas, although not statistically significant.

Selection bias cannot be excluded as the study is hospital-based. Malignant epithelial cancers were a priori excluded from controls as experimental evidence suggests that a similar association with vitamin-A may hold for epithelial neoplasms other than lung cancer. A further subset of diseases could have been related to modification of diet, and the possible selection bias which might have occurred was investigated by comparing these groups of diagnoses (namely groups 4 to 6 in Table 1) with the others. No differences were found in the consumption of carrots and green vegetables after adjustment for the relevant confounders. Other determinants of a differential selection of cases and controls by dietary habits could be taken into account (places of birth and residence, and nutrition); tobacco smoking, which proved to be a minor confounder, was accurately controlled for by stratification or regression. Several sources of information bias should be considered. Interviewer bias can be reasonably excluded as interviewers were aware of the case-control status of the subjects but not of the alimentary hypothesis; standardization by interviewer did not result in appreciable changes of the estimates.

As vitamin-A intake was not the chief objective of the investigation, the food items included in the questionnaire were necessarily limited to those considered as the main sources of retinol (liver and cheese) and carotenoids (carrots and green vegetables) in the Italian diet; a further limitation is that only recent diet was investigated. Moreover information collected (frequencies of consumption) was such that only a vague exposure categorization could be applied. The well-known effect of an unbiased misclassification on the estimate of relative risk is to reduce it towards one. Misclassification of exposure is more likely to have occurred for frequencies of carrot and vegetable intake since these are often cooked with other foodstuffs while liver and cheese constitute well-defined dishes which are more easily recalled. It seems unlikely therefore that the observed difference between dietary sources of retinol and carotene, is due to misclassification.

Concerning histological subtypes of lung cancer, a cohort and two case-control studies found a protective effect of vitamin-A and carotene for epidermoid and small cell carcinomas, but not for adenocarcinomas. In our data the protective effect associated with carrot consumption was evident for all subtypes, although not significant for adenocarcinomas.

Since no systematic review of pathological diagnoses has been performed, misclassification could affect this result.

Given the limitations discussed above our data support the evidence of a protective effect of vegetables, but while it is possible that the nutrient responsible is carotene, the contribution of other substances present in the foods must be considered—possibly other vitamin compounds.

Moreover disparate results obtained for the two foodstuffs by morphology and by smoking status, suggest different protective mechanisms. In particular a further hypothesis to be considered is that prevention might intervene at later stages of lung carcinogenesis as suggested by having observed the association only among current smokers. Several dietary case-control studies on lung cancer observed higher relative risks among heavy smokers, but a recent report observed an inverse association only among ex-smokers.

**ACKNOWLEDGEMENTS**

The collection of data was supported by the US National Cancer Institute under contract N01-CP-05644. The analysis was supported by the Italian National Research Council (Progetto Finalizzato Oncologia) under contract no. 84.00448.44.

The authors wish to thank the staff of the hospitals in the Varese province and in Milan who collaborated during data collection, and Professor A Morabito,
Institute of Biometry, University of Milan, for advice on statistical analysis.

REFERENCES


(Revised version received January 1986)