

Published in final edited form as:

Ann Epidemiol. 2011 July ; 21(7): 543–550. doi:10.1016/j.annepidem.2010.11.019.

Principal component analysis of dietary and lifestyle patterns in relation to risk of subtypes of esophageal and gastric cancer

Stephanie A. Navarro Silvera, Susan T Mayne, Harvey A. Risch, Marilie D Gammon, Thomas Vaughan, Wong-Ho Chow, Joel A Dubin, Robert Dubrow, Janet Schoenberg, Janet L Stanford, A. Brian West, Heidrun Rotterdam, and William J Blot

Montclair State University, Department of Health and Nutrition Sciences, Montclair, NJ (Stephanie A.N. Silvera). Yale University School of Public Health, Division of Chronic Disease Epidemiology, New Haven CT (Susan T. Mayne, Harvey A. Risch, Robert Dubrow), Yale School of Medicine, Department of Pathology, New Haven CT (A. Brian West). University of North Carolina School of Public Health, Department of Epidemiology, Chapel Hill, NC (Marilie D. Gammon). Fred Hutchinson Cancer Research Center, Program in Epidemiology, and University of Washington, School of Public Health and Community Medicine, Department of Epidemiology, Seattle, WA (Thomas Vaughan and Janet L. Stanford). National Cancer Institute, Division of Cancer Epidemiology and Genetics, NIH, DHHS, Bethesda, MD (Wong-Ho Chow). New Jersey Department of Health and Senior Services, Cancer Epidemiology Services, Trenton NJ (Janet Schoenberg, deceased). Columbia University, Department of Pathology, New York, NY (Heidrun Rotterdam). International Epidemiology Institute, Rockville, MD. and Vanderbilt University, Nashville, TN (William J Blot). University of Waterloo, Department of Statistics & Actuarial Science, and Department of Health Studies and Gerontology, Waterloo, Ontario (Joel A. Dubin).

Abstract

Purpose—To perform pattern analyses of dietary and lifestyle factors in relation to risk of esophageal and gastric cancers.

Methods—We evaluated risk factors for esophageal adenocarcinoma (EA), esophageal squamous cell carcinoma (ESCC), gastric cardia adenocarcinoma (GCA), and other gastric cancers (OGA) using data from a population-based case-control study conducted in Connecticut, New Jersey, and western Washington state. Dietary/lifestyle patterns were created using principal component analysis (PCA). Impact of the resultant scores on cancer risk was estimated through logistic regression.

Results—PCA identified six patterns: meat/nitrite, fruit/vegetable, smoking/alcohol, legume/meat alternate, GERD/BMI, and fish/vitamin C. Risk of each cancer under study increased with rising meat/nitrite score. Risk of EA increased with increasing GERD/BMI score, and risk of ESCC rose with increasing smoking/alcohol score and decreasing GERD/BMI score. Fruit/vegetable scores were inversely associated with EA, ESCC, and GCA.

Conclusions—PCA may provide a useful approach for summarizing extensive dietary/lifestyle data into fewer interpretable combinations that discriminate between cancer cases and controls. The analyses suggest that meat/nitrite intake is associated with elevated risk of each cancer under

© 2011 Elsevier Inc. All rights reserved.

Please address correspondence and reprint requests to Stephanie A.N. Silvera, Montclair State University, Department of Health and Nutrition Sciences, 1 Normal Ave, Montclair, NJ 07043. Telephone 973-655-2125; FAX 973-655-5461 silveras@mail.montclair.edu..

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

study, while fruit/vegetable intake reduces risk of EA, ESCC, and GCA. GERD/obesity were confirmed as risk factors for EA and smoking/alcohol as risk factors for ESCC.

Keywords

esophageal adenocarcinoma; gastric cardia adenocarcinoma; esophageal squamous cell carcinoma; diet; principal components

Introduction

An increasing incidence of adenocarcinomas of the esophagus (EA) and the gastric cardia (GCA), though to a lesser extent, has been well noted in the literature, with incidence rising over 6-fold between 1973 and 2002 (1-3). The increasing trends are persisting in the U.S. (4) and have been found in Great Britain, Australia, The Netherlands, Denmark and other western nations (5). In response to the rising trend, the study described herein was initiated in the United States in order to investigate potential risk factors. This, and other, studies have thus far shown that obesity (6-9), tobacco use (10-13), and gastroesophageal reflux disease (GERD) (14-16) are important risk factors for EA and GCA, and that *Helicobacter pylori* colonization (17-19) may be an important protective factor.

It has been theorized that fruits and vegetables, which are high in antioxidants, phytosterols and other substances, may inhibit carcinogenesis by free-radical quenching or by blocking the formation of *N*-nitroso compounds (20-22). While epidemiologic studies have found that fruits and vegetables are associated with a decreased risk of gastric and esophageal cancers without regard to subsite or histologic type (23, 24), evidence linking dietary factors to subtypes of these cancers is more limited. In earlier reports of the present multi-center, population-based study, we found significant inverse associations between intake of nutrients found primarily in plant-based foods and the risk of EA and GCA (25, 26). In another U.S. population-based study, Brown et al. (6) observed a significantly reduced risk of EA among white men reporting the highest intake of raw fruits, raw vegetables and cruciferous vegetables. While they did not find a consistent association between consumption of meat, poultry, and fish and risk of EA (6), we observed a significant positive association between intake of meat and animal protein and risk of EA and GCA (25, 26). We and others have also presented evidence of an inverse association between dietary fiber intake and risk of EA (6, 25) and GCA (25, 27). For non-cardia gastric adenocarcinoma (OGA), the available evidence suggests a positive association with nitrite-containing foods in western countries and salted or preserved foods in Asian countries (24).

Analyses of independent effects of individual dietary items are complicated by the correlations in consumption of foods in the typical diet within and between food groups. For example, intakes of fruits and vegetables are associated positively and negatively with consumption of other food groups (28). Also, patterns of dietary intake are correlated with other factors known to affect health such as smoking and SES (29-32). Thus, studies that focus on individual nutrients or food groups may overlook these correlations as well as potential interactions among differing foods or food groups. Alternative analyses of diet and lifestyle patterns have therefore begun to be explored to try to help explain the role of diet in chronic disease etiology (33-35). Principal components analysis (PCA) makes it possible to identify potentially interpretable patterns in the data by weighting variables within a principal component. The resulting weighted linear combinations of variables in a principal component identify variables that co-vary and lend themselves to interpretations and assignments of titles that have common sense meanings. The use of an analysis of this nature allows for the capture of variation in overall food intake among the study subjects in a nicely interpretable manner. Four such analyses have examined esophageal (36, 37) and

gastric cancer (36-39) outcomes. The findings of these studies are discussed in detail, in light of our results, in the discussion section.

The analyses presented herein supplement our previous nutrient and food group analyses (25, 26) based on more traditional epidemiologic methods of dietary assessment/analysis by using PCA to examine patterns of these same dietary factors along with other established risk factors for these cancers (e.g., body mass index (BMI), GERD, smoking).

Subjects and methods

Study population

The methods of subject recruitment and data collection have been previously described in detail (12). A multi-center, population-based, case-control study of adenocarcinoma of the esophagus, squamous cell carcinoma of the esophagus (ESCC), adenocarcinoma of the gastric cardia, and adenocarcinoma of other anatomic sites of the stomach was conducted in three geographic areas of the U.S. with population-based tumor registries: the entire state of Connecticut, a 15-county area of New Jersey, and a three-county area of western Washington State. The project sought to enroll four population-based case groups of approximately equal size, namely patients aged 30-79, newly diagnosed in 1993-95, with one of the four types of cancer mentioned above, along with a population-based control group. Institutional review board approval was obtained from all participating centers and from the Connecticut Department of Public Health. Certain data used in this study were obtained from the Connecticut Tumor Registry, located in the Connecticut Department of Public Health. The authors assume full responsibility for analyses and interpretation of these data.

Attempts were made to recruit all subjects diagnosed with EA and GCA (target cases). A random sample of subjects diagnosed with ESCC and OGA were selected to serve as comparison cases and were frequency-matched to target cases by five-year age group, sex and geographic site. Cases were identified via rapid reporting systems in each of the three areas, with pathology reports sought for all potentially eligible cases. Two study pathologists systematically reviewed slides and medical records of potential cases to determine final eligibility.

Controls, randomly selected from the general population of the study areas, were frequency matched to the expected distribution of target cases by five-year age group and sex. Waksberg's random digit dialing method was utilized in order to identify controls aged 30-64 (40); while those who were aged 65-79 years of age were identified by stratified random sampling of Health Care Financing Administration rosters.

Complete interviews were obtained for 80.6% of eligible target subjects (cases of EA and GCA), 74.1% of comparison case subjects (cases of ESCC and OGA), and 70.2% of eligible controls, with a mean time between diagnosis and case interview of 3.7 months. A total of 1,839 individuals were interviewed for this study. Of these, 34 subjects were seriously ill and unable to complete the dietary portion of the questionnaire and were therefore excluded from the analyses. An additional 23 persons were excluded from analysis due to implausible energy intake (< 600 Kcal/d, n = 20 or >5000 Kcal/d, n = 3). The dietary analyses thus included interviews of 1,782 subjects: 687 controls, 282 cases with EA, 255 with GCA, 206 with ESCC, and 352 with OGA. Proxy interviews were more common among cases (EA = 31%, GCA = 26%, ESCC = 35%, and OGA = 30%) than among controls (3.4%).

Data collection

After obtaining written informed consent, trained interviewers administered a structured in person questionnaire to the study participant, or if necessary, to a close relative who served as a proxy respondent. The questionnaire contained questions on demographics including height and usual adult weight (used to calculate BMI), tobacco and alcohol, other beverage use (e.g., coffee, tea), medical history including frequency and duration of GERD symptoms, use of medications, and occupational history. A food frequency questionnaire developed and previously validated (41) by investigators at the Fred Hutchinson Cancer Research Center (FHCRC) was adapted to assess usual consumption in the period 3-5 years before diagnosis (cases) or interview (controls). Subjects were asked to report how often they consumed 104 different foods. Data from the food frequency questionnaire were entered and verified, then sent to the FHCRC for processing, and were initially linked with the University of Minnesota, Nutrition Coding Center Nutrient Data system for estimation of nutrient intake. Average daily intake of nitrite was estimated separately, through software and databases developed by the authors (42).

Statistical analysis

Food group and subgroup variables were defined as previously described (26). Twenty-eight diet and lifestyle variables were selected for evaluation due to their associations with either esophageal or gastric cancers, either in these data or in other published studies, and included in the initial principal component analysis. Principal component analysis was used for variable reduction in the face of redundant, correlated exposures, so that a more meaningful logistic regression may be performed, the results from which will have greater interpretability. Diet and lifestyle principal components (i.e., specific linear combinations of the variables studied) were generated using the SAS software principal components procedure (proc factor method=prin; version 8.2; SAS Institute, Inc, Cary, NC) using data from controls only according to the approach outlined by both Hatcher (43) and Timm (44). The following variables were included in the principal component analysis procedure: 19 food groups, usual adult body mass index (BMI), average number of cigarettes smoked per day, consumption of beer, wine and liquor (drinks per day, each separately), intake of fiber (g/day), vitamin C (mg/day) and nitrite (mg/day), and reported frequency of GERD symptoms (ordered categorical variable, 6 levels).

Principal components with observed eigenvalues greater than 1.0 in the principal component analysis were further evaluated. The larger the absolute value of a loading for a variable to a principal component, the greater the contribution of that variable to that principal component. Variables that had loadings of 0.20 or greater were considered to be making, at minimum, a reasonable contribution to the principal component; however, all loadings (those above and below 0.20) were included in calculating principal components scores. In addition, the second principal component is independent of the first principal component and the third is independent of the first two principal components, and so on. Labeling of the principal components was done based upon our interpretation of the data.

After a varimax rotation, which typically makes the principal component more interpretable, principal component scores were generated using the SAS Proc Factor command (43, 44). These scores represent the weighted sums of the exposure variables, where the weights are equal to the principal component loading. The scores were then categorized into quartiles and used in multivariate unconditional logistic regression analyses for each cancer subtype to calculate odds ratios (OR) and corresponding 95% confidence intervals (CI). All models were mutually adjusted for all other principal components from the factor-loading matrix and were also adjusted for covariates of interest, including study site (Connecticut/Washington/New Jersey), age, gender, race (white/other), proxy interview status (proxy/

non-proxy), income (ordered categorical variable, 6 levels), education (ordered categorical variable, 7 levels), and energy intake. All tests of significance were two-sided, with a *p* value of 0.05 considered statistically significant.

Results

Principal component loadings for each of the variables of interest based on data from the controls are shown in Table 1. Six patterns were retained in the analysis, which accounted for 48% of the total variance. The first pattern loaded heavily on nitrite, high-nitrite meats, and red meats, and was therefore termed a Meat/Nitrite pattern. In addition, high-fat dairy, vitamin C, refined grains, fiber, poultry and starchy vegetables also loaded with this pattern. In contrast, the second principal component, termed a Fruit/Vegetable pattern, loaded heavily on raw, deep yellow/orange and dark green and cruciferous vegetables, tomato products, as well as both citrus and non-citrus fruits. The third principal component was distinguished from the others as it loaded most heavily on alcohol and cigarette use, and was thus termed a Smoking/Alcohol pattern. The fourth pattern was characterized by its heavy loading on legumes and meat alternates, and was therefore termed a Legume/Meat Alternate pattern. The fifth principal component loaded heavily on GERD symptoms and usual adult BMI, and was termed a GERD/BMI pattern. Whole grains and low-fat dairy also loaded in this pattern. The sixth and final pattern meeting our predetermined criteria loaded most heavily on fish and dietary vitamin C intake. This was termed a Fish/Vitamin C pattern.

Esophageal cancer

For EA risk, significant inverse associations were found with highest versus lowest quartile of the fruit/vegetable principal component ($OR_{Q4 \text{ vs. } Q1} = 0.43$, 95% CI: 0.26, 0.71; $P_{Trend} < 0.001$), whereas significant positive associations were found with the highest versus lowest quartile of meat/nitrite principal component ($OR_{Q4 \text{ vs. } Q1} = 5.61$, 95% CI: 2.81, 11.20; $P_{Trend} < 0.001$) and the GERD/BMI principal component ($OR_{Q4 \text{ vs. } Q1} = 3.50$, 95% CI: 2.01, 6.09; $P_{Trend} < 0.001$), shown in Table 2.

For ESCC risk, a strong and significant positive association was found with the smoking/alcohol principal component ($OR_{Q4 \text{ vs. } Q1} = 10.82$, 95% CI: 5.16, 22.68; $P_{Trend} < 0.001$) (Table 2). A positive association was also found with the meat/nitrite principal component ($OR_{Q4 \text{ vs. } Q1} = 2.01$, 95% CI: 0.82, 4.95; $P_{Trend} = 0.06$) and the fish/vitamin C principal component ($OR_{Q4 \text{ vs. } Q1} = 1.83$, 95% CI: 0.99, 3.37; $P_{Trend} = 0.05$). In contrast, significant inverse associations were found with the fruit/vegetable principal component ($OR_{Q4 \text{ vs. } Q1} = 0.36$, 95% CI: 0.19, 0.69; $P_{Trend} = 0.001$) and the GERD/BMI pattern ($OR_{Q4 \text{ vs. } Q1} = 0.40$, 95% CI: 0.21, 0.77; $P_{Trend} < 0.001$).

Gastric cancer

A positive association was observed for the meat/nitrite principal component and GCA risk ($OR_{Q4 \text{ vs. } Q1} = 1.82$, 95% CI: 0.91, 3.65; $P_{Trend} = 0.04$), while a negative association was observed for the fruit/vegetable score ($OR_{Q4 \text{ vs. } Q1} = 0.63$, 95% CI: 0.39, 1.01; $P_{Trend} = 0.04$) (Table 3). The meat/nitrite pattern was significantly associated with increased risk of OGA ($OR_{Q4 \text{ vs. } Q1} = 2.40$, 95% CI: 1.25, 4.62; $P_{Trend} = 0.01$, Table 3).

Discussion

In this relatively large population-based case-control study of men and women in the U.S. we identified six dietary/lifestyle patterns, which subsequently were shown to tend to discriminate cases from controls. The meat/nitrite principal component was consistently positively associated with each of the four cancers of interest, with the strongest association

seen for EA. Significant positive associations were also found between the GERD/BMI pattern and EA risk, and between the smoking/alcohol pattern and ESCC risk. In contrast, significant inverse associations were found between the fruit/vegetable pattern and risk of both types of esophageal cancer. These findings underscore and build upon our previous reports using these data (7, 12, 25, 26).

The elevated risk of EA associated with the meat/nitrite pattern is consistent the findings of Bahmanyar and Ye (37) who reported a 60% increased risk with a high-meat “Western diet” pattern in an analysis of dietary patterns and EA. In addition, Chen et al. (36) reported a 3.6-fold elevated risk of EA with a high-meat dietary pattern and Campbell et al. (38), in an analysis of dietary patterns and gastric adenocarcinomas, reported a statistically significant positive association with a Western dietary pattern among men.

Thus the patterns we observed with the meat/nitrite principal component tend to fit well with previous observations. This principal component loaded high for red meats, high-nitrite meats, and nitrite, but also for high-fat dairy products, poultry, refined grains, dietary fiber, vitamin C, and starchy vegetables, so that inferences cannot be made regarding meats or any specific food product. Also, while the literature generally indicates that fiber is inversely associated with EGA (24), in our study population, some main sources of fiber were refined grains (white bread, rice, crackers, sugar cereals, pancakes, pizza, and pasta), which are not particularly fiber-rich foods but were consumed in larger quantities than whole grain foods. The case-control difference with respect to this principal component was particularly strong for EA, with a greater than 5-fold increased risk between those with high versus low scoring on the principal component. This principal component may be indicative of a generalized western diet, high in red meats, starchy vegetables and refined grains, of note because EA seems to be a cancer of western nations and not yet increasingly evident in developing or eastern countries.

Our finding that higher fruit/vegetable principal component scores were significantly negatively associated with both subtypes of esophageal cancer adds to the literature showing an inverse association of fruit and vegetable intake with EA and ESCC (6, 10, 25-27, 36, 45-50). Bahmanyar and Ye (37), in contrast, found no association between their “Healthy Diet” dietary pattern, which was characterized by high fruit, vegetable, fish and poultry consumption, and risk of EA.

In contrast to other investigations (24), we found no association between the fruit/vegetable principal component and OGA, but a borderline association with GCA. Campbell et al. (38), in keeping with our finding of a suggested inverse association between the fruit/vegetable principal component and GCA, reported an inverse association between their “Prudent” dietary pattern, which correlated with higher fruit and vegetable consumption, and risk of GCA among both men and women.

In our analyses, the GERD/BMI principal component pattern was associated with a 3.5-fold increased risk of EA. Chronic reflux has emerged along with obesity as an important risk factor for EA (3). In contrast to the association with EA, the GERD/BMI pattern was significantly inversely associated with ESCC. This case group differed both from controls and the three other case groups in a number of ways, including being significantly leaner and consuming fewer calories per day. At least 3 large studies have found an inverse association between ESCC and BMI (51-53). The inverse association between the GERD/BMI pattern and ESCC was stronger prior to adjustment for energy intake (data not shown).

We observed a significant 10-fold increased risk of ESCC in the highest versus lowest quartile of the smoking/alcohol principal component, a finding expected since smoking and drinking are known to be the major risk factors for ESCC (2). Bahmanyar and Ye (37)

likewise reported a 3.5-fold increased risk of ESCC associated with their “High alcohol” pattern, which was characterized by high beer and liquor intake.

As with case-control studies generally, the present study has several strengths and limitations, including the potential for recall bias. The tumor-type specificity of risks, however, which, for example, included an inverse association between intake of fruits and vegetables and esophageal cancers and GCA, but not OGA, argues against this bias to some extent. We also repeated the PCA on the combined cases and controls and the same components resulted from the analysis, with the exception that the third and fourth components arose in reverse order. That is, the smoking/alcohol PC was third among controls only and fourth among the combined cases and controls. Ultimately, this did not materially alter the results from the logistic regression analysis and therefore, would not change the study conclusions. In addition, due to the high case-fatality rate of these cancers, direct interview data could not be obtained from approximately 30% of cases. When separate analyses excluded proxy interviews, however, the odds ratio estimates remained virtually unchanged.

While our data are consistent with the literature indicating that consuming a diet high in a combination of fruits and vegetables is inversely associated with risk of both subtypes of esophageal cancer and with GCA, they also indicated that smoking and drinking did not load strongly with any diet patterns. It is therefore unlikely that dietary findings, particularly for ESCC, are due to residual confounding, which is often a concern. Also, meat intake did not load strongly with the fruit and vegetable principal component, suggesting that these intakes may have opposing effects on cancer risk. In addition, we found that high-nitrite meat and red meats, loaded together, along with starchy foods, refined grains, and high fat dairy. The resulting principal component, similar to unhealthy “Western” diets reported by other studies (37, 38) suggests that future analyses of nitrite might need to adjust for non-nitrite rich meats, refined and starchy foods, and possibly high fat dairy products. Further, like our previous reports (7, 15) these alternative analyses implicate both GERD and obesity as key risk factors for EA. Thus, principal component analysis, which allows for the examination of underlying correlations between variables of interest, particularly dietary and lifestyle components, may yield useful practical conclusions and recommendations for future research, targeted prevention, and framing specific health behavior messages.

Acknowledgments

We thank the following: study managers Ms. Sarah Greene and Ms. Linda Lannom (Westat), data management Ms. Shelley Niwa (Westat), and field supervisors Mrs. Patricia Owen (Connecticut), Mr. Tom English (New Jersey), and Ms. Berta Nicol-Blades (Washington) for data collection and processing; Dr. Alan Kristal for assistance in designing and processing the dietary questionnaires; the Yale Cancer Center Rapid Case Ascertainment Shared Resource, the 178 hospitals in Connecticut, New Jersey, and Washington for their participation in the study; and the study participants. This research was supported [in part] by the Intramural Research Program of the NIH, National Cancer Institute, Division of Cancer Epidemiology and Genetics.

Funding United States Public Health Service (U01-CA57983, U01-CA57949, U01-CA57923, P30ES10126); National Cancer Institute, National Institutes of Health, Department of Health and Human Services (N02-CP40501, N01-CN05230).

References

1. Blot WJ, Devesa SS, Kneller RW, Fraumeni JF Jr. Rising incidence of adenocarcinoma of the esophagus and gastric cardia. *JAMA*. 1991; 265:1287–1289. [PubMed: 1995976]
2. Blot, WJ.; McLaughlin, JK.; Fraumeni, JF., Jr., editors. *Esophageal Cancer*. Oxford University Press; Oxford: 2006.

3. Holmes RS, Vaughan TL. Epidemiology and pathogenesis of esophageal cancer. *Semin Radiat Oncol.* 2007; 17:2–9. [PubMed: 17185192]
4. Brown LM, Devesa SS, Chow WH. Incidence of adenocarcinoma of the esophagus among white Americans by sex, stage, and age. *J Natl Cancer Inst.* 2008; 100 Epub.
5. Bollschweiler E, Wolfgarten E, Gutschow C, Holscher AH. Demographic variations in the rising incidence of esophageal adenocarcinoma in white males. *Cancer.* 2001; 3:549–555. [PubMed: 11505399]
6. Brown LM, Swanson CA, Gridley G, Swanson GM, Schoenberg JB, Greenberg RS, et al. Adenocarcinoma of the esophagus: Role of obesity and diet. *J Natl Cancer Inst.* 1995; 87:104–109. [PubMed: 7707381]
7. Chow WH, Blot WJ, Vaughan TL, Risch HA, Gammon MD, Stanford JL, et al. Body mass index and risk of adenocarcinomas of the esophagus and gastric cardia. *J Natl Cancer Inst.* 1998; 90:150–155. [PubMed: 9450576]
8. Lagergren J, Bergtrom R, Nyren O. Association between body mass and adenocarcinoma of the esophagus and gastric cardia. *Ann Intern Med.* 1999; 130:883–890. [PubMed: 10375336]
9. Vaughan TL, Davis S, Kristal A, Thomas DB. Obesity, alcohol, and tobacco as risk factors for cancers of the esophagus and gastric cardia: Adenocarcinoma versus squamous cell carcinoma. *Cancer Epidemiol Biomark Prev.* 1995; 4:85–92.
10. Anderson LA, Watson RGP, Murphy SJ, Johnston BT, Comber H, Mc Guigan J, et al. Risk factors for Barrett's oesophagus and oesophageal adenocarcinoma: results from the FINBAR study. *World J Gastroenterol.* 2007; 13:1585–1594. [PubMed: 17461453]
11. Chow WH, Swanson CA, Lissowska J, Groves FD, Sobin LH, Nsierowska-Guttmejer A, et al. Risk of stomach cancer in relation to consumption of cigarettes, alcohol, tea and coffee in Warsaw, Poland. *Int J Cancer.* 1999; 81:871–876. [PubMed: 10362132]
12. Gammon MD, Schoenberg JB, Ahsan H, Risch HA, Vaughan TL, Chow WH, et al. Tobacco, alcohol, and socioeconomic status and adenocarcinomas of the esophagus and gastric cardia. *J Natl Cancer Inst.* 1997; 89:1277–1284. [PubMed: 9293918]
13. Lagergren J, Bergtrom R, Lindgren A, Nyren O. The role of tobacco, snuff and alcohol use in the aetiology of cancer of the oesophagus and gastric cardia. *Int J Cancer.* 2000; 85:340–346. [PubMed: 10652424]
14. Chow WH, Finkle W, McLaughlin JK, Frankl H, Ziel HK, Fraumeni JF Jr. The relation of gastroesophageal reflux disease and its treatment to adenocarcinomas of the esophagus and gastric cardia. *JAMA.* 1995; 274:474–477. [PubMed: 7629956]
15. Farrow DC, Vaughan TL, Sweeney C, Gammon M, Chow WH, Risch HA, et al. Gastroesophageal reflux disease, use of H2 receptor antagonists, and risk of esophageal and gastric cancer. *Cancer Causes Control.* 2000; 11:231–238. [PubMed: 10782657]
16. Lagergren J, Bergtrom R, Lindgren A, Nyren O. Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma. *New England J Med.* 1999; 340:825–831. [PubMed: 10080844]
17. Wong A, Fitzgerald RC. Epidemiologic risk factors for Barrett's esophagus and associated adenocarcinoma. *Clin Gastroenterol Hepatol.* 2005; 3:1–10. [PubMed: 15645398]
18. Chow WH, Blaser MJ, Blot WJ, Gammon MD, Vaughan TL, Risch HA, et al. An inverse relation between cagA+ strains of *Helicobacter pylori* infection and risk of esophageal and gastric cardia adenocarcinoma. *Cancer Res.* 1998; 58:588–590. [PubMed: 9485003]
19. Kamangar F, Dawsey SM, Blaser MJ, Perez-Perez GI, Pietinen P, Newschaffer CJ, et al. Opposing risks of gastric cardia and noncardia gastric adenocarcinomas associated with *Helicobacter pylori* seropositivity. *J Natl Cancer Inst.* 2006; 98:1445–1452. [PubMed: 17047193]
20. Mirvish SS, Wallcave L, Eagen M, Shubik P. Ascorbate-nitrite reaction: possible means of blocking the formation of carcinogenic N-nitroso compounds. *Science.* 1972; 177:65–68. [PubMed: 5041776]
21. Raineri R, Weisburger JH. Reduction of gastric carcinogens with ascorbic acid. *Ann N Y Acad Sci.* 1975; 258:181–189. [PubMed: 1060401]
22. Steinmetz KA, Potter JD. Vegetables, fruit, and cancer. II. Mechanisms. *Cancer Causes Control.* 1991; 2:427–442. [PubMed: 1764568]

23. Vainio, H.; Bianchini, F. IARC Handbooks of cancer prevention. Vol. Volume 8: fruit and vegetables. IARC Press, International Agency for Research on Cancer; Lyon (France): 2003.
24. World Cancer Research Fund. Food, physical activity and the prevention of cancer: a global perspective. American Institute for Cancer Research; Washington DC: 2007.
25. Mayne ST, Risch R, Dubrow R, Chow WH, Gammon MD, Vaughan T, et al. Nutrient intake and risk of subtypes of esophageal and gastric cancer. *Cancer Epidemiol Biomark Prev.* 2001; 10:1055–1062.
26. Silvera, SA Navarro; Mayne, ST.; Risch, HA.; Gammon, MD.; Vaughan, TL.; Chow, WH., et al. Food group intake and risk of subtypes of esophageal and gastric cancer. *Int J Cancer.* 2008; 123:852–860. [PubMed: 18537156]
27. Terry P, Lagergren J, Ye W, Wolk A, Nyren O. Inverse association between intake of cereal fiber and risk of gastric cardia cancer. *Gastroenterol.* 2001; 120:387–391.
28. Gordon T, Fisher M, Rifkind BM. Some difficulties inherent in the interpretation of dietary data from free-living populations. *Am J Clin Nutr.* 1984; 39:152–156. [PubMed: 6606975]
29. Kant AK, Schatzkin A, Graubard BI, Schairer C. A prospective study of diet quality and mortality in women. *JAMA.* 2000; 283:2109–2115. [PubMed: 10791502]
30. La Vecchia C, Munoz SE, Braga C, Fernandez E, Decarli A. Diet diversity and gastric cancer. *Int J Cancer.* 1997; 72:255–257. [PubMed: 9219829]
31. Randall E, Marshal JR, Graham S, Brasure J. High-risk health behaviors associated with various dietary patterns. *Nutr Cancer.* 1991; 16:135–151. [PubMed: 1796009]
32. Wirfalt AK, Jeffery RW. Using cluster analysis to examine dietary patterns: nutrient intakes, gender, and weight status differ across food pattern clusters. *J Am Diet Assoc.* 1997; 97:272–279. [PubMed: 9060944]
33. McCullough ML, Feskanich D, Stampfer MJ, Rosner BA, Hu FB, Hunter DJ, et al. Adherence to the Dietary Guidelines for Americans and risk of major chronic disease in women. *Am J Clin Nutr.* 2000; 72:1214–1222. [PubMed: 11063452]
34. McCullough ML, Feskanich D, Rimm EB, Giovannucci EL, Ascherio A, Variyam JN, et al. Adherence to the Dietary Guidelines for Americans and risk of major chronic disease in men. *Am J Clin Nutr.* 2000; 72:1223–1231. [PubMed: 11063453]
35. Slattery ML, Boucher KM, Caan B, Potter J, Ma KN. Eating patterns and risk of colon cancer. *Am J Epidemiol.* 1998; 148:4–16. [PubMed: 9663397]
36. Chen H, Ward MH, Graubard BI, Heineman EF, Markin RM, Potischman NA, et al. Dietary patterns and adenocarcinoma of the esophagus and distal stomach. *Am J Clin Nutr.* 2002; 75:137–144. [PubMed: 11756071]
37. Bahmanyar S, Ye W. Dietary patterns and risk of squamous-cell carcinoma and adenocarcinoma of the esophagus and adenocarcinoma of the gastric cardia: A population-based case-control study in Sweden. *Nutr Cancer.* 2006; 54:171–178. [PubMed: 16898861]
38. Campbell PT, Sloan M, Kreiger N. Dietary patterns and risk of incident gastric cancer. *Am J Epidemiol.* 2007; 167:295–304. [PubMed: 18048377]
39. Palli D, Russo A, Decarli A. Dietary patterns, nutrient intake and gastric cancer in a high-risk area of Italy. *Cancer Causes Control.* 2001; 12:163–172. [PubMed: 11246845]
40. Waksberg J. Sampling methods for random digit dialing. *J Am Stat Assoc.* 1978; 73:40–46.
41. Kristal AR, Feng Z, Coates FJ, Oberman A, George V. Associations of race, ethnicity, education and dietary intervention on the validity and reliability of a food frequency questionnaire in the Women's Health Trial Feasibility Study in Minority Populations. *Am J Epidemiol.* 1997; 146:856–869. [PubMed: 9384206]
42. Risch HA, Jain M, Choi NW, Fodor JG, Pfeiffer CJ, How GR, et al. Dietary factors and the incidence of cancer of the stomach. *Am J Epidemiol.* 1985; 122:947–959. [PubMed: 2998182]
43. Hatcher, L. A step-by-step approach to using the SAS System for factor analysis and structural equation modeling. SAS Institute Inc.; Cary, NC: 1994.
44. Timm, NH. Applied Multivariate Analysis. Springer; New York: 2002.

45. Bosetti C, La Vecchia C, Talamini R, Simonato L, Zambon P, Negri E, et al. Food groups and risk of squamous cell esophageal cancer in northern Italy. *Int J Cancer*. 2000; 87:289–294. [PubMed: 10861489]
46. Brown LM, Swanson CA, Gridley G, Swanson GM, Silverman DT, Greenberg RS, et al. Dietary factors and risk of squamous cell esophageal cancer among black and white men in the United States. *Cancer Causes Control*. 1998; 9:467–474. [PubMed: 9934713]
47. Castellsague X, Munoz N, De Stefani E, Victoria CG, Castelletto R, Rolon PA. Influence of mate drinking, hot beverages and diet on esophageal cancer risk in South America. *Int J Cancer*. 2000; 88:658–664. [PubMed: 11058886]
48. De Stefani E, Brennan P, Boffetta P, Ronco A, Mendilaharsu M, Deneo-Pellegrini H. Vegetables, fruits, related dietary antioxidants, and risk of squamous cell carcinoma of the esophagus: a case-control study in Uruguay. *Nutr Cancer*. 2000; 38:23–29. [PubMed: 11341040]
49. Gao YT, McLaughlin JK, Gridley G, Blot WJ, Ji BT, Dai Q, et al. Risk factors for esophageal cancer in Shanghai, China. II. Role of diet and nutrients. *Int J Cancer*. 1994; 58:197–202. [PubMed: 8026881]
50. Tzonou A, Lipworth L, Garidou A, Signorello LB, Lagiou P, Hsieh C, et al. Diet and risk of esophageal cancer by histologic type in a low-risk population. *Int J Cancer*. 1996; 68:300–304. [PubMed: 8903470]
51. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*. 2008; 371:569–578. [PubMed: 18280327]
52. Smith M, Zhou M, Whitlock G, Yang G, Offer A, Hui G, et al. Esophageal cancer and body mass index: results from a prospective study of 220,000 men in China and a meta-analysis of published studies. *Int J Cancer*. 2008; 122:1604–1610. [PubMed: 18059032]
53. Engeland A, Tretli S, Bjorge T. Height and body mass index in relation to esophageal cancer; 23-year follow-up of two million Norwegian men and women. *Cancer Causes & Control*. 2004; 15:837–843. [PubMed: 15456997]

Table 1

Assessment of factor-loading matrix of lifestyle principal components from controls of a United States multi-center, population-based case-control study of esophageal and gastric cancer (1993-1995)

Variables	Meat/ Nitrite	Fruit/ Vegetable	Smoking/ Alcohol	Legume/ Meat Alternate	GERD/ BMI	Fish/ Vitamin C
Juice		0.26				
Citrus fruits		0.55			0.20	
Non-citrus fruits		0.56		0.21	0.32	
Cruciferous vegetables		0.49		0.32		0.25
Deep yellow & orange vegetables		0.59				
Dark green vegetables		0.67			-0.20	
Starchy vegetables	0.42			0.25		
Raw vegetables		0.82				
Legumes				0.73		
Tomato products	0.25	0.67				
High-fat dairy	0.49			0.22	0.27	
Low-fat dairy					0.42	0.30
Whole grains				0.22	0.55	
Refined grains	0.50					
Fish		0.21				0.81
Poultry	0.50					0.31
High-nitrite meats	0.77					
Red meats	0.67					
Meat alternates	0.23			0.73		
Dietary fiber	0.84			0.20		
Dietary Vitamin C	0.67					0.50
Nitrite	0.86					
GERD symptoms/ wk				-0.24	0.41	
Usual BMI					0.46	
Cigarettes/day		-0.22	0.46			-0.32
Beer			0.67			

Variables	Meat/ Nitrite	Fruit/ Vegetable	Smoking/ Alcohol	Legume/ Meat Alternate	GERD/ BMI	Fish/ Vitamin C
Wine			<i>0.58</i>			
Liquor			<i>0.78</i>			
Proportion of Variance	21%	14%	6%	6%	5%	5%

• Values of <0.20 have been excluded from the table for simplicity, loadings for non-displayed variables were included in the analyses.

• Values with loading of 0.40 or greater are bolded and italicized.

Table 2

Adjusted logistic regression – Lifestyle factors and risk of esophageal cancer, from United States multicenter, population-based study (1993-1995). Results are mutually adjusted^a OR and 95% CI, comparing principal components scores (quartiles, 1 = low) and cancer risk.

		EA											
Quartile	Meat/Nitrite OR	Fruit/Vegetable		Smoking/Alcohol		Legume/Meat Alternate		GERD/BMI		Fish/ Vitamin C			
		OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI		
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00		
2	1.48	0.91, 2.42	0.64	0.41, 1.01	1.15	0.73, 1.81	0.71	0.45, 1.12	2.62	1.55, 4.43	1.00	0.64, 1.56	
3	3.19	1.82, 5.59	0.58	0.37, 0.92	1.19	0.74, 1.90	0.91	0.58, 1.44	1.86	1.08, 3.21	0.72	0.45, 1.15	
4	5.61	2.81, 11.20	0.43	0.26, 0.71	0.89	0.54, 1.46	0.98	0.60, 1.60	3.50	2.01, 6.09	0.94	0.59, 1.48	
P for Trend	<0.001	<0.001		0.74			0.88		<0.001			0.48	

		ESCC											
Quartile	Meat/Nitrite OR	Fruit/Vegetable		Smoking/Alcohol		Legume/Meat Alternate		GERD/BMI		Fish/ Vitamin C			
		OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI		
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00		
2	1.28	0.69, 2.38	0.88	0.51, 1.54	3.72	1.78, 7.77	1.06	0.59, 1.90	0.47	0.27, 0.81	1.41	0.78, 2.57	
3	2.28	1.13, 4.61	0.63	0.35, 1.12	5.91	2.93, 12.37	0.98	0.53, 1.78	0.34	0.19, 0.61	1.57	0.87, 2.84	
4	2.01	0.82, 4.95	0.36	0.19, 0.69	10.82	5.16, 22.68	1.09	0.57, 2.07	0.40	0.21, 0.77	1.83	0.99, 3.37	
P for Trend	0.06	0.001		<0.001			0.88		<0.001			0.05	

^a Adjusted for gender, age, site, race, income, education, proxy status, energy intake, and mutual adjustment for other principle components.

Table 3

Adjusted logistic regression – Lifestyle factors and risk of gastric cancer, from United States multicenter, population-based study. Results are mutually adjusted^a OR and 95% CI, comparing principal components scores (quartiles, 1 = low) and cancer risk.

GCA												
Quartile	Meat/Nitrite		Fruit/Vegetable		Smoking/Alcohol		Legume/Meat Alternate		GERD/BMI		Fish/ Vitamin C	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
1	1.00		1.00		1.00		1.00		1.00		1.00	
2	1.13	0.70, 1.82	0.92	0.58, 1.43	1.18	0.75, 1.85	0.80	0.50, 1.26	1.27	0.79, 2.02	1.07	0.68, 1.68
3	1.75	1.03, 2.96	0.79	0.50, 1.26	1.24	0.78, 1.97	0.90	0.56, 1.44	0.84	0.52, 1.36	1.06	0.67, 1.65
4	1.82	0.91, 3.65	0.63	0.39, 1.01	0.84	0.51, 1.39	0.96	0.59, 1.56	1.14	0.68, 1.89	1.11	0.70, 1.74
P for Trend	0.04		0.04		0.63		0.98		0.94		0.68	

OGA												
Quartile	Meat/Nitrite		Fruit/Vegetable		Smoking/Alcohol		Legume/Meat Alternate		GERD/BMI		Fish/ Vitamin C	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
1	1.00		1.00		1.00		1.00		1.00		1.00	
2	1.89	1.23, 2.92	1.10	0.71, 1.70	1.14	0.76, 1.73	0.79	0.52, 1.20	1.00	0.65, 1.53	1.15	0.74, 1.78
3	2.03	1.23, 3.35	0.98	0.63, 1.52	1.24	0.80, 1.90	0.89	0.58, 1.36	1.05	0.70, 1.59	1.16	0.76, 1.78
4	2.40	1.25, 4.62	0.94	0.60, 1.46	0.84	0.53, 1.33	0.73	0.46, 1.15	0.77	0.48, 1.24	1.31	0.86, 1.99
P for Trend	0.01		0.64		0.62		0.27		0.40		0.22	

^a Adjusted for gender, age, site, race, income, education, proxy status, energy intake, and mutual adjustment for other principle components.