

# Dietary Patterns and Prostate Cancer Risk in the National Health and Nutrition Examination Survey Epidemiological Follow-up Study Cohort

**Marilyn Tseng** Division of Population Science, Fox Chase Cancer Center, Philadelphia, Pennsylvania

**Rosalind A. Breslow** National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland

**Robert F. DeVellis** Department of Health Behavior and Health Education, University of North Carolina, Chapel Hill, North Carolina

**Regina G. Ziegler** Epidemiology and Biostatistics Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland.

## Abstract

Ecological studies implicate a “Western” diet in prostate cancer development, but whether dietary patterns measured in individuals are associated with risk has not been studied previously. We examined this issue using prospective data from the nationally representative United States Health Examination Epidemiological Follow-up Study. Among 3,779 men followed from 1982–84 to 1992, 136 incident cases were identified. Using principal component analysis on responses to a 105-item dietary questionnaire, the following three distinct patterns were identified: a vegetable-fruit pattern; a red meat-starch pattern characterized by red meats, potatoes, cheese, salty snacks, and desserts; and a Southern pattern characterized by such foods as cornbread, grits, sweet potatoes, okra, beans, and rice. In adjusted proportional hazards models, prostate cancer risk was not associated with the vegetable-fruit or red meat-starch pattern, but higher intake of the Southern pattern showed a reduction in risk (3rd versus 1st tertile relative risk, 0.6; 95% confidence interval, 0.4–1.1; trend  $P=0.08$ ) that approached statistical significance. The inverse association was observed in black and non-black men and was not attributable to intake of any individual foods or nutrients. A Southern dietary pattern may reflect a history of living in the South and serve as an integrative marker of sunlight exposure and protection through 1,25-dihydroxyvitamin D production. Further evaluation and better characterization of the pattern would offer more information on potentially beneficial features of the diet or its associated lifestyle.

## Introduction

In 2003, prostate cancer is expected to remain the most commonly diagnosed cancer and the second leading cause of cancer death in United States men (1). Although risk factors for prostate cancer have been identified, only age, ethnicity, and family history of prostate cancer are well-established (1). Thus, feasible measures for primary prevention of the disease remain limited.

Epidemiological and laboratory research suggests that dietary intake is one factor that might be modified to reduce risk (2, 3). In particular, ecological (4–6) and migrant (7) studies have implicated a “Western” dietary pattern as a risk factor for prostate cancer. Whether dietary patterns measured at the level of the individual are associated with prostate cancer risk has not been studied previously, in part because of the relative novelty of the approach but also because standard methods for identifying, measuring, and interpreting dietary patterns are only now being developed. By examining exposure to several related dietary factors simultaneously, quantifying the aggregate risk associated with a particular combination of foods, and offering results that are based on actual dietary practice and more easily translated into useful recommendations, a dietary pattern approach provides a useful complement to findings based on single nutrients or single food groups.

In research using principal components analysis (PCA) to identify and quantify dietary patterns, two patterns emerge fairly consistently in samples from the United States: one characterized by intake of vegetables and fruits and a Western-style pattern based on red meat and starch (8–10). The Western pattern has been related to increased risk of colon cancer (8, 11), cardiovascular disease (12), and diabetes (13). The objectives of our analyses were to identify dietary patterns in a nationally representative sample of United States men using PCA and to examine for the first time their associations with prostate cancer risk in prospectively collected data, with the goal of clarifying the importance of specific dietary patterns to the development of prostate cancer.

## Materials and Methods

**Study Population.** The study sample included male participants in the first National Health and Nutrition Examination Survey (NHANES) I Epidemiological Follow-up Study (NHEFS). NHANES I, conducted between 1971 and 1975, used a multistage sampling design to obtain a national probability sample of the noninstitutionalized civilian population of the United States, excluding Alaska, Hawaii, and Native American reservation lands (14, 15). The elderly and persons residing in poverty areas were oversampled. About 70% of those sampled were both interviewed and medically examined in NHANES I.

NHEFS was a longitudinal study of the 14,407 participants between the ages of 25 and 74 years at the time of the initial survey (16–19). Participants were followed for health and vital status through 1992. At interviews conducted in 1982–84, 1986, 1987, and 1992, participants or their proxies were interviewed. Also, health records were obtained for instances in which participants reported an overnight stay in a health care facility between the baseline examination and last follow-up visit. Death certificates were obtained for deaths during the follow-up period and were identified by the National Death Index or other tracing mechanisms. Health records were obtained for over 70% of reported overnight stays, and death certificates were obtained for 99% of deaths occurring between 1971–75 and the 1992 follow-up (19).

Because food frequency data obtained in the 1971–75 interview included questions on only 13 broad food categories, we used more detailed data from a 105-item food frequency questionnaire administered in 1982–84. Thus, 1982–84 served as the baseline for these analyses. Of the 14,407 NHEFS participants, 5,811 were men. Of these, 1,202 died before the 1982–84 interview, 351 could not be traced, and 333 were not interviewed in 1982–84. Subjects were further excluded from the remaining study sample if they had a diagnosis of prostate cancer at or before the 1982–84 interview ( $n=57$ ), did not complete the diet questionnaire ( $n=79$ ), or had energy intake of  $<500$  or  $>4,400$  kcal/day ( $n=10$ ), leaving 3,779 men available for analysis.

**Identification of Prostate Cancer Cases.** Cases of invasive prostate cancer were identified following a procedure described by Breslow *et al.* (20). Briefly, potential cases were all men with an International Classification of Diseases, Ninth Revision, Clinical Modification code of 185 (invasive prostate cancer), 233.4 (prostate carcinoma *in situ*), V10.46 (personal history of malignant prostate neoplasm), or 60.3–60.5 (prostatectomy surgical procedures) recorded in at least one of the following sources of data: (a) a first diagnosis of prostate cancer reported at any of the follow-up interviews conducted in 1986, 1987, or 1992; (b) one or more hospital stays during the follow-up period with a discharge diagnosis coded as any of the codes given above; (c) a death certificate with underlying or nonunderlying cause of death coded as any of the codes given above. Archived records of interviews and overnight health care facility stays were then reviewed. No *in situ* cases were considered in our analyses. “Definite” case status was assigned if cancer could be confirmed from histopathology reports or medical records. Determinations based only on interview or death certificate data were assigned “probable” case status. Of 136 cases diagnosed during the follow-up of the 3,779 men from the 1982–84 interview, 89 were considered “definite” cases, and an additional 47 were considered “probable” cases.

**Data Collection.** Information on dietary intake was obtained from a 105-item food frequency questionnaire administered in the 1982–84 interview. The questionnaire was designed to include foods commonly consumed in the United States diet and covered the major food groups, including meats, fish, poultry, grains, fruits, vegetables, dairy products, sweets, snacks, and beverages. Intake of specific nutrients such as energy, total fat, and vitamin A was estimated by multiplying frequency of intake of each food by the nutrient content for the food’s portion size. Because the 1982–84 NHEFS dietary interview collected only frequency information, information on nutrient content and portion size for each food was based on sex- and age-specific 24-h recall data from the NHANES II, a separate national survey conducted in 1976–80. A detailed description of the method used to assign nutrient content and portion size to each food item in the NHEFS dietary questionnaire using NHANES II data has been published (21).

Other information available from the 1982–84 interview included race, place of residence, longest held occupation, family income, first-degree family history of prostate cancer, current weight, alcohol intake, smoking behavior, sun exposure, level of physical activity, and current multivitamin use. Information on height and level of education was available from the 1971–75 interview.

**Identification of Dietary Patterns.** Patterns of food intake were identified by PCA (22, 23) using frequency responses to the dietary questionnaire. (An example of SAS programming statements used to run the analysis is provided at <http://www.fccc.edu/research/labs/tseng/TsengDOD01.html>.) Individuals were randomly placed into one of two equally sized groups, or split-samples, to confirm reproducibility of the principal components identified. For the first split sample, a matrix of correlations among frequency of consumption for the questionnaire food items was

constructed and entered in the PCA. Extraction of principal components was followed by orthogonal rotation of retained components to allow for interpretability (22, 23). The number of components to retain for rotation was based on examination of scree plots and interpretability of the components (23); although another common strategy is to rotate all factors with eigenvalues  $>1.0$ , this method has been shown to overestimate the number of components (23). The analysis was repeated in the second split sample to confirm reproducibility of results. Cronbach's coefficient  $\alpha$  (24) was used to evaluate internal consistency for each component retained. In psychometric research, a coefficient  $\alpha$  of  $\geq 0.70$  generally indicates acceptable reliability (25), although in previous research, dietary pattern scales with coefficient  $\alpha$  as low as 0.5–0.6 were predictive of disease (26). As an additional assessment of the robustness of the patterns identified, we used oblique rather than orthogonal rotation, but the same patterns emerged.

A component score was calculated for each dietary pattern for each individual to represent the individual's level of intake for the pattern. The score for each pattern was computed as a linear composite of the foods with meaningful loadings ( $\geq 0.20$ ) for only that pattern. Scores were calculated by taking the unweighted sum of standardized frequencies of intake for each food associated with the pattern. When we computed pattern scores as a linear composite of all variables weighted based on regression results (27), scores that were calculated the two different ways were highly correlated ( $r > 0.85$ ), and estimates of relative risk (RR) for prostate cancer were similar.

We examined construct validity of the patterns, or the extent to which they behave as expected theoretically with respect to other variables (28), by describing their associations with sociodemographic and lifestyle variables among 3,544 men with complete variable data. The variables, selected based on social and historical descriptions of the development of those patterns (29–31), included age, place of residence (rural, urban, suburban), socioeconomic status (SES), and various health-related behaviors.

**Data Analysis.** Active follow-up ended in 1992, but non-cases were censored at the last date they were known to be alive and free of prostate cancer: specifically, at their date of last interview. Follow-up time was calculated by subtracting the 1982–84 interview date from the censoring date for non-cases and the interview date from the date of prostate cancer diagnosis for cases. For four cases identified from death certificate data only, the 1982–84 interview date was subtracted from date of death rather than from date of diagnosis.

Adjusted RR of prostate cancer was estimated for tertiles of pattern scores using Cox proportional hazards models while adjusting for age (continuous years) and race (white, black, or other race). Other variables including United States region, urban/rural residence, education, first-degree family history of prostate cancer, current body mass index, recreational physical activity, recreational and occupational sun exposure, multivitamin use, smoking status, and past and current alcohol consumption were evaluated as confounders based on their associations with predictor and response variables and by comparing unadjusted and adjusted estimates from regression analyses. Final multivariate models included 3,616 men with complete covariate data and were adjusted for age, race, United States region (Northeast, Midwest, South, West), residence (rural, urban, suburban), education ( $<$ high school, high school completion,  $>$ high school), recreational sun exposure (little, occasional, frequent), recreational physical activity (little/none, moderate, much), smoking status (never, former, current), current alcohol intake (none, little, moderate, heavy), and energy intake (tertiles). All covariates were coded using dummy variables to allow for nonlinear associations across categories. Controlling for energy as a continuous variable produced no meaningful changes in estimates. Although proxy interview data were obtained for 2.5% of participants in the 1982–84 NHEFS interview, none of the dietary data for the 3,616 men included in our final analyses was obtained through a proxy.

*P*s for linear trend were obtained for each dietary pattern by including an ordinal variable representing the scaled median value for each tertile in the multivariate model controlling for the covariates listed above. To examine the possibility of effect modification by race, we ran proportional hazards models in black and non-black men separately; men of ethnicities other than white or black were too few ( $n=46$ ) to allow for separate analysis. Because of the small number of black cases ( $n=27$ ), we dichotomized pattern scores for all men at the median value for black men. *P*s for interaction were obtained from a model including all men, with a pattern category  $\times$  race interaction term.

In multivariate models controlling for the same covariates, we also examined the effects of specific foods and nutrients potentially related to risk of prostate cancer, including red meat, dairy, fruits and vegetables, tomatoes, energy, total and saturated fat, calcium, vitamin A, and dietary fiber (32–35). Nutrient values were log-transformed as necessary and energy-adjusted using the residual method (36). RRs were estimated for tertiles of intake relative to the lowest tertile, but for infrequently consumed items such as okra and grits, estimates were for consumption *versus* nonconsumption. To account for sample weighting from the survey design, all final models also included the following design variables: age ( $<65$  *versus*  $\geq 65$  years), poverty census enumeration district (residence *versus* nonresidence), and family income ( $< \$3,000$ ,  $\$3,000$ – $\$6,999$ ,  $\$7,000$ – $\$9,999$ ,  $\$10,000$ – $\$14,999$ , and  $\geq \$15,000$ ; Ref. 37), although results from models with and without design variables were similar.

## Results

Mean age of the men in the study sample was 58 years, 11% were African-American, and their usual residence was roughly equally distributed among the four regions of the United States. Over a mean follow-up of 7.6 years (range <1–10.7 years), 136 prostate cancer cases were identified in the cohort of 3,779 men.

In PCA, the following three dietary patterns emerged consistently across the split samples (Table 1): (a) a “vegetable-fruit” pattern with high loadings for vegetables, fruits, fish, and shellfish; (b) a “red meat-starch” pattern with high loadings for red meats, potatoes, salty snacks, cheese, sweets, and desserts; and (c) a “Southern” pattern with high loadings for beans, rice, and such traditionally Southern United States foods as corn-bread, grits, sweet potatoes, and okra. The same three patterns emerged when we conducted the analysis in black men only (results not shown). Thus, calculation of pattern scores was based on the PCA solution including all men.

As an assessment of construct validity, we described the distributions of selected sociodemographic and health-related characteristics across pattern tertiles (Table 2). Men with high intake of the vegetable-fruit pattern were more likely to be white and of higher SES, to live in the Northeastern and Western United States and in suburbs, and to use multivitamins, exercise, and not smoke. Men with high intake of the red meat-starch pattern were more likely to be white and younger in age, to live in rural areas and in the Midwest, and to smoke, exercise, and not use multivitamins. Associations of the red meat-starch pattern with SES indicators reflected neither especially high nor especially low SES; *e.g.*, men with high intake were more likely to have graduated high school but not college and were more likely to be above the poverty level but not at the highest incomes. In contrast, men with high intake of the Southern pattern were more likely to be black, to be of lower SES, to live in rural areas and in the South, and to smoke and not use multivitamins.

In Cox proportional hazards models (Table 3), although the red meat-starch pattern was not associated with prostate cancer, intermediate intake of the vegetable-fruit pattern showed a slight elevation of risk [RR, 1.5; 95% confidence interval (CI), 0.9–2.3]. However, for neither pattern was there evidence of a trend, and none of the point estimates reached statistical significance. Higher intake of the Southern pattern showed a trend ( $P=0.08$ ) and a reduction in risk in the third tertile (3rd *versus* 1st tertile RR, 0.6; 95% CI, 0.4–1.1) that approached statistical significance. Risk estimates were not materially different when we excluded 14 prostate cancer cases diagnosed within a year of the dietary interview, reclassified 47 “probable” cases as non-cases, conducted analyses using SUDAAN (38), or used age rather than time-on-study as the time scale (39). In race-specific analyses, the association between the Southern dietary pattern and prostate cancer risk was more pronounced in black men (above *versus* below median RR, 0.2; 95% CI, 0.2–0.6), but estimates were based on only 27 cases among black men, and the  $P$  for interaction ( $P=0.47$ ) was not significant.

We explored possible explanations for the inverse association for the Southern pattern by examining each of the foods and nutrients associated with it, but none was associated with prostate cancer risk (results not shown). Dietary patterns were also correlated with intake of other specific foods and nutrients that have previously been linked to prostate cancer risk (Table 2), but with the exception of dairy foods and calcium (described in a separate study), none of these were clearly associated with disease in our data; RR estimates for red meat, tomatoes, energy, total and saturated fat, fiber, and vitamins A and C were all around one and showed no trend in either direction (results not shown). However, men with an intermediate level of fruit intake had a nonsignificantly elevated risk (2nd *versus* 1st tertile RR, 1.3; 95% CI, 0.8–2.1).

Since 1986, when the United States Food and Drug Administration approved the prostate-specific antigen test for monitoring prostate cancer progression, incidence has increased more steeply in men of higher SES and, presumably, with better awareness of or access to screening modalities (40). To explore the possibility of detection bias, we conducted additional analyses including only cases identified before 1986. The inverse association for the Southern pattern persisted (3rd *versus* 1st tertile RR, 0.4; 95% CI, 0.2–0.9), but estimates were based on only 46 cases.

## Discussion

In a nationally representative sample of men, we identified the following three dietary patterns: a vegetable-fruit pattern, a red meat-starch pattern, and a Southern pattern. The red meat-starch pattern was not associated with disease, but intermediate intake of the vegetable-fruit pattern was nonsignificantly associated with increased prostate cancer risk. Intake of the Southern pattern showed a trend that was suggestive of an inverse association and that could not be attributed to any specific foods within the pattern.

Identification of the vegetable-fruit and Western patterns in this sample is consistent with findings of previous United States studies (8–10) and with anthropological and historical accounts of traditional American eating habits (29–31). Moreover, their associations with sociodemographic and health-related characteristics were as expected based on observations of the historical emergence of those patterns (29), confirming the validity of their measurement using PCA (28,41). Although other studies have identified patterns specific to Mexican Americans (26, 42), ours is the first, to our

knowledge, to identify a Southern United States pattern in a sample not limited to a specific ethnic or regional group. Our results further suggest that the pattern is not a spurious finding; *e.g.*, the pattern emerged across split samples in our analyses, was easily recognizable as a distinct pattern, and its associations with socio-demographic characteristics were consistent with social/cultural descriptions of the pattern (31), supporting the pattern's construct validity.

We surmise that we were able to identify the Southern pattern because the food frequency questionnaire included such specifically Southern items as cornbread, grits, and okra, and because we did not group the 105 food items from the questionnaire for the analysis. We chose not to collapse food items for several reasons. First, creating groups of potentially dissimilar foods may diminish the ability to identify more specific patterns. Indeed, when we collapsed foods into 35 predefined food groups (9), only two rather than three patterns clearly emerged: the vegetable-fruit and red meat-starch patterns (results not shown). Ability to identify dietary patterns, therefore, is strongly dependent both on the food items included in the instrument and on how foods are aggregated into groups for analysis. That the Southern pattern was reproducible across split samples and was associated with other variables as theoretically expected suggests that its identification in our data was not a spurious finding; rather, collapsing foods into groups might have prevented finding a true pattern.

Grouping foods before PCA may also attenuate or increase the variance of measures of association between dietary patterns and disease (43). Collapsing foods into groups would likely have produced patterns that explained more of the total variation in food intake than the 11% explained by the three patterns in our study. However, the primary objective of performing PCA in diet-disease studies is not to explain total variation but to examine associations of conceptually meaningful patterns with disease risk. Indeed, McCann *et al.* (43) have demonstrated that increasing the amount of variance explained by collapsing foods into groups does not improve estimates of disease risk.

Factors that might have obscured or biased associations in our study merit discussion. Using PCA to quantify dietary patterns may involve some measurement error, for example. However, reasonably high (>0.60) coefficient  $\alpha$  for the three patterns indicates good internal reproducibility for each pattern, and using an alternative method to calculate pattern scores (27) produced similar associations with risk. Also worth noting with respect to measurement error is that, in contrast to indices in which uncorrelated items are selected based on *a priori* criteria, PCA-based scores represent the common and presumably causal source of variation underlying the item set; by aggregating information in correlated items, PCA essentially maximizes signal (shared source of variation) and minimizes noise (item-specific sources of variation) in the sum scores.

Associations may also be obscured if undiagnosed cases with early-stage tumors were included among non-cases. Although prostate-specific antigen testing was relatively uncommon before 1991 (44), some bias in detecting prostate cancer cases remains possible as well, given the associations of dietary patterns with urban/rural residence and SES. Information on access to screening and screening behavior was not available to evaluate this possibility directly. Excluding *in situ* cancers in this analysis, however, likely minimized bias that might result from including early-stage, PSA-identified tumors. In addition, RR estimates were largely unchanged when we controlled for sociodemographic factors that may be linked to screening such as education (40) and when we limited cases to those identified before government approval of prostate-specific antigen testing in 1986.

We found a slightly elevated prostate cancer risk with intermediate intake of the vegetable-fruit pattern but no clear trend. When we examined fruits and vegetables separately, we observed no association for vegetables but a slightly elevated risk for an intermediate level of fruit intake. Our finding is similar to that of other studies that observed a positive association with fruit intake (32), but the explanations for this finding are not known. We observed no elevation in risk for selected nutrients associated with fruit or vegetable intake.

Our results do not support the hypothesis that a Western pattern increases risk of prostate cancer. In our sample, red meat-starch pattern intake, intake of red meat as a food group, and intake of energy, total fat, and saturated fat were not associated with disease, although previous cohort studies have fairly consistently found positive associations for red meat and for saturated and animal fat (34). Besides detection bias, this may also reflect simply the lack of strong influence of overall adult diet on risk, insufficient variability in intake, or inaccurate measurement of the underlying pattern. A clearer understanding of dietary pattern measurement is warranted before more definite conclusions can be drawn. Notably, a Western pattern was also not associated with colorectal or breast cancer in recent analyses in the Swedish Mammography cohort (45, 46).

We observed a nonsignificant but suggestive inverse association for the Southern pattern. This finding is especially intriguing because black men were more likely to consume this pattern but remained at higher risk for prostate cancer. In race-specific analyses, the apparent inverse association persisted in both black and non-black men. The association was not attributable to any individual foods within the pattern or to any nutrients of prior interest. Our finding suggests that prostate cancer incidence might increase with movement away from a traditional Southern cuisine. Alternatively,

the finding raises additional questions regarding interpretation of dietary pattern measures. For example, a Southern dietary pattern may reflect a history of living in the South and serve as an integrative marker of sunlight exposure rather than a simple measure of overall dietary habits. Sunlight has been hypothesized to protect against prostate cancer by catalyzing synthesis of 7-dehydro-cholesterol in the skin to 25-hydroxyvitamin D, which is subsequently converted in the kidney, and possibly in the prostate, to its biologically active form, 1,25-dihydroxyvitamin D (47, 48). In experimental studies, 1,25-dihydroxyvitamin D reduces cell proliferation, induces cell differentiation and apoptosis, and disturbs cell survival signals in the signaling pathway (49, 50); in animal models, these effects have been confirmed in prostate epithelial and cancer cells (51). Ecological analyses in the United States have demonstrated a North-South trend in prostate cancer mortality, with lower rates in the South (52), and a recent epidemiological study also offers evidence that both childhood and cumulative lifetime sun exposure is associated with reduced risk (53). Although we controlled for other life-style factors in our analysis, including current residence in the South, and we considered several others as potential confounders in preliminary analyses, it remains possible that our measure for the Southern pattern represents earlier-life or long-term sunlight exposure. Our findings emphasize the importance of considering the context of any given dietary pattern to understand the relevance of the diet or of its associated lifestyle on health status.

In summary, we found that prostate cancer risk was not clearly associated with either the red meat-starch or the vegetable-fruit pattern, but we observed a suggestive, inverse association for the Southern pattern. The association was observed in both black and non-black men and was not attributable to any individual foods within the pattern or to any nutrients of prior interest. A Southern dietary pattern may reflect a history of living in the South and serve as an integrative marker of sunlight exposure and protection through 1,25-dihydroxyvitamin production. However, better characterization of the pattern would offer more information on potentially beneficial features of the diet or of its associated lifestyle. Thus, our findings should be explored and confirmed in other data to clarify interpretation of these observations. Although a pattern approach might yield a valuable perspective in diet-disease studies, strategies for improving methods of identifying and evaluating dietary patterns also require additional consideration.

### Acknowledgments

We thank Marianne Hyer for assistance in bringing together the relevant data, Dr. Barry Graubard for assistance on the analysis, and Dr. W. Thomas London for input on the manuscript.

### References

1. American Cancer Society. Cancer Facts & Figures 2003. Atlanta, GA: American Cancer Society, 2003.
2. Clinton, S. K., and Giovannucci, E. Diet, nutrition, and prostate cancer. *Annu. Rev. Nutr.*, 18: 413–440, 1998.
3. Chan, J. M., Stampfer, M. J., and Giovannucci, E. L. What causes prostate cancer? A brief summary of the epidemiology. *Semin. Cancer Biol.*, 8: 263–273, 1998.
4. Rose, D., Boyar, A., and Wynder, E. International comparisons of mortality rates for cancer of the breast, ovary, prostate, and colon and per capita food consumption. *Cancer (Phila.)*, 58: 2363–2371, 1986.
5. Yu, H., Harris, R. E., Gao, Y. T., Gao, R., and Wynder, E. L. Comparative epidemiology of cancers of the colon, rectum, prostate and breast in Shanghai, China versus the United States. *Int. J. Epidemiol.*, 20: 76–81, 1991.
6. Tominaga, S., and Koroishi, T. An ecological study on diet/nutrition and cancer in Japan. *Int. J. Cancer, Suppl* 10: 2–6, 1997.
7. Kolonel, L., Hankin, J., and Nomura, A. Multiethnic studies of diet, nutrition, and cancer in Hawaii. *In: Hayashi, Y., Nagao, M., Sugimura, T., Takayama, S., Tomatis, L., Wattenberg, L.-W., and Wogan, G. N. (eds.), Diet, Nutrition, and Cancer*, pp. 29–40. Tokyo: VNU Sci. Press, 1986.
8. Slattery, M. L., Boucher, K. M., Caan, B. J., Potter, J. D., and Ma, K. N. Eating patterns and risk of colon cancer. *Am. J. Epidemiol.*, 148: 4–16, 1998.
9. Hu, F., Rimm, E., Smith-Warner, S., Feskanich, D., Stampfer, M., Ascherio, A., Sampson, L., and Willett, W. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am. J. Clin. Nutr.*, 69: 243–249, 1999.
10. Tseng, M., and DeVellis, R. F. Fundamental dietary patterns and their correlates among US whites. *J. Am. Diet. Assoc.*, 101: 929–932, 2001.
11. Randall, E., Marshall, J. R., Brasure, J., and Graham, S. Dietary patterns and colon cancer in western New York. *Nutr.*

Cancer, 18: 265–276, 1992.

12. Hu, F. B., Rimm, E. B., Stampfer, M. J., Ashcerio, A., Spiegelman, D., and Willett, W. C. Prospective study of major dietary patterns and risk of coronary heart disease in men. *Am. J. Clin. Nutr.*, 72: 912–921, 2000.
13. van Dam, R. M., Rimm, E. B., Willett, W. C., Stampfer, M. J., and Hu, F. B. Dietary patterns and risk for type 2 diabetes mellitus in U. S. men. *Ann. Intern. Med.*, 136: 201–209, 2002.
14. Miller, H. W. Plan and Operation of the Health and Nutrition Examination Survey. United States–1971–73. *Vital Health Stat. 1, No. 10a*. U.S. Department of Health and Human Services: Hyattsville, MD, 1973.
15. National Center for Health Statistics. Plan and Operation of the Health and Nutrition Examination Survey. United States–1971–73. *Vital Health Stat. 1, No. 10b*. U.S. Department of Health and Human Services: Hyattsville, MD, 1977.
16. Cohen, B. B., Barbano, H. E., Cox, C. S., Feldman, J. J., Finucane, F. F., Kleinman, J. C., and Madans, J. H. Plan and Operation of the NHANES I Epidemiologic Followup Study, 1982–84. *Vital Health Stat. 1, No. 22*. U.S. Department of Health and Human Services: Hyattsville, MD, 1987.
17. Finucane, F. F., Freid, V. M., Madans, J. H., Cox, C. S., Kleinman, J. C., Rothwell, S., Barbano, H. E., and Feldman, J. C. Plan and operation of the NHANES I Epidemiologic Followup Study, 1986. *Vital Health Stat. 1, No. 25*. U.S. Department of Health and Human Services: Hyattsville, MD, 1990.
18. Cox, C., Rothwell, S., Madans, J., Finucane, F. F., Freid, V. M., Kleinman, J. C., Barbano, H. E., and Feldman, J. J. Plan and operation of the NHANES I Epidemiologic Followup Study, 1987. *Vital Health Stat. 1, No. 27*. U.S. Department of Health and Human Services: Hyattsville, MD, 1992.
19. Cox, C. S., Mussolino, M. E., Rothwell, S. T., Lane, M. A., Golden, C. D., Madans, J. H., and Feldman, J. J. Plan and operation of the NHANES I Epidemiologic Follow-up Study 1992. *Vital Health Stat. 1, No. 35*. U.S. Department of Health and Human Services: Hyattsville, MD, 1997.
20. Breslow, R. A., Wideroff, L., Graubard, B. I., Erwin, D., Reichman, M. E., Ziegler, R. G., and Ballard-Barbash, R. Alcohol and prostate cancer in the NHANES I Epidemiologic Follow-up Study. *Ann. Epidemiol.*, 9: 254–261, 1999.
21. Ursin, G., Ziegler, R. G., Subar, A. F., Graubard, B. I., Haile, R. W., and Hoover, R. Dietary patterns associated with a low-fat diet in the National Health Examination Follow-up Study. *Am. J. Epidemiol.*, 137: 916–927, 1993. 22. Kleinbaum, D. G., Kupper, L. L., and Muller, K. E. *Applied regression analysis and other multivariable methods*. 2nd edition, Belmont, CA: Duxbury Press, 1988.
23. Kline, P. *An Easy Guide to Factor Analysis*. New York: Routledge, 1994.
24. Cronbach, L. J. Coefficient  $\alpha$  and the internal structure of tests. *Psychometrika*, 16: 297–334, 1951.
25. Nunnally, J. *Psychometric Theory*. New York: McGraw-Hill, 1978.
26. Tseng, M., DeVellis, R. F., Kohlmeier, L., Khare, M., Maurer, K. R., Everhart, J. E., and Sandler, R. S. Patterns of food intake and gallbladder disease in Mexican Americans. *Public Health Nutr.*, 3: 233–243, 2000.
27. Hatcher, L. *A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling*. Cary, NC: SAS Institute, Inc., 1994.
28. DeVellis, R. F. *Scale Development Theory and Applications*. Applied Social Research Methods Series. Vol. 26, Newbury Park, CA: SAGE Publications, 1991.
29. Levenstein, H. A. *Revolution at the Table: The Transformation of the American Diet*. New York: Oxford University Press, 1988.
30. McIntosh, E. N. *American Food Habits in Historical Perspective*. Westport, CT: Praeger Publishers, 1995.
31. Pillsbury, R. *No Foreign Food: The American Diet in Time and Place*. Boulder, CO: Westview Press, 1998.
32. Chan, J. M., and Giovannucci, E. L. Vegetables, fruits, associated micronutrients, and risk of prostate cancer. *Epidemiol. Rev.*, 23: 82–86, 2001.
33. Chan, J. M., and Giovannucci, E. L. Dairy products, calcium, and vitamin D and risk of prostate cancer. *Epidemiol. Rev.*, 23: 87–92, 2001.
34. Kolonel, L. N. Fat, meat, and prostate cancer. *Epidemiol. Rev.*, 23: 72–81, 2001.
35. Giovannucci, E. Tomatoes, tomato-based products, lycopene, and cancer: review of the epidemiologic literature. *J. Natl. Cancer Inst. (Bethesda)*, 91: 317–331, 1999.
36. Willett, W., and Stampfer, M. Total energy intake: implications for epidemiologic analyses. *Am. J. Epidemiol.*, 124: 17–27, 1986.
37. Korn, E. L., and Graubard, B. I. Epidemiologic studies utilizing surveys: accounting for the sampling design. *Am. J. Public Health*, 81: 1166–1173, 1991.

38. Shah, B. V., Barnwell, B. G., and Bieler, G. S. SUDAAN User's Manual, Release 7.0. Research Triangle Park, NC: Research Triangle Institute, 1996.
39. Korn, E. L., Graubard, B. I., and Midthune, D. Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. *Am. J. Epidemiol.*, *145*: 72–80, 1997.
40. Liu, L., Cozen, W., Bernstein, L., Ross, R. K., and Deapen, D. Changing relationship between socioeconomic status and prostate cancer incidence. *J. Natl. Cancer Inst. (Bethesda)*, *98*: 705–709, 2001.
41. Last, J. M. (ed.). *A Dictionary of Epidemiology*, 2nd edition. Oxford Uni-
42. Wolff, C. B., and Wolff, H. K. Maternal eating patterns and birth weight of Mexican American infants. *Nutr. Health*, *10*: 121–134, 1995.
43. McCann, S. E., Marshall, J. R., Brasure, J. R., Graham, S., and Freudenheim, J. L. Analysis of patterns of food intake in nutritional epidemiology: food classification in principal components analysis and the subsequent impact on estimates for endometrial cancer. *Public Health Nutr.*, *4*: 989–997, 2001.
44. Etzioni, R., Penson, D. F., Legler, J. M., di Tommaso, D., Boer, R., Gann, P. H., and Feuer, E. J. Overdiagnosis due to prostate-specific antigen screening: Lessons from U. S. prostate cancer incidence trends. *J. Natl. Cancer Inst. (Bethesda)*, *94*: 981–990, 2002.
45. Terry, P., Hu, F. B., Hansen, H., and Wolk, A. Prospective study of major dietary patterns and colorectal cancer risk in women. *Am. J. Epidemiol.*, *154*: 1143–1149, 2001.
46. Terry, P., Suzuki, R., Hu, F. B., and Wolk, A. A prospective study of major dietary patterns and the risk of breast cancer. *Cancer Epidemiol. Biomark. Prev.*, *10*: 1281–1285, 2001.
47. Schwartz, G. G., and Hulka, B. S. Is vitamin D deficiency a risk factor for prostate cancer? *Anticancer Res.*, *10*: 1307–1312, 1990.
48. Holick, M. F. Calcium and vitamin D. Diagnostics and therapeutics. *Clin. Lab. Med.*, *20*: 569–590, 2000.
49. Miller, G. J. Vitamin D and prostate cancer: biologic interactions and clinical potentials. *Cancer Metastasis Rev.*, *17*: 353–360, 1999.
50. Johnson, C. S., Hershberger, P. A., and Trump, D. L. Vitamin D-related therapies in prostate cancer. *Cancer Metastasis Rev.*, *21*: 147–158, 2002.
51. Giovannucci, E. L. Dietary influences of 1,25(OH)<sub>2</sub> vitamin D in relation to prostate cancer: a hypothesis. *Cancer Causes Control*, *9*: 567–582, 1998.
52. Hanchette, C. L., and Schwartz, G. G. Geographic patterns of prostate cancer mortality. Evidence for a protective effect of ultraviolet radiation. *Cancer (Phila.)*, *15*: 2861–2869, 1992.
53. Luscombe, C. J., Fryer, A. A., French, M. E., Liu, S., Saxby, M. F., Jones, P. W., and Strange, R. C. Exposure to ultraviolet radiation: association with susceptibility and age at presentation with prostate cancer. *Lancet*, *358*: 641–642, 2001.

Table 1 Factor loadings for foods associated with each dietary pattern, in split samples of 3,779 adult male participants in the National Health Examination Follow-up Study, 1982–84

	Vegetable-fruit		Red meat-starch		Southern			
	Sample 1 <sup>a</sup>	Sample 2 <sup>b</sup>	Sample 1	Sample 2	Sample 1	Sample 2		
Carrots	55	44	Salty snacks <sup>c</sup>	47	35	Combread or hush puppies	60	58
Iceberg/head lettuce	49	47	Fried potatoes	44	39			
Broccoli	48	49	Pork <sup>d</sup>	42	22	Okra	51	65
Cauliflower	46	47	Mixed dishes with meat or cheese	42	35	Sweet potatoes or yellow yams	48	56
Leaf lettuce	45	38						
Salad dressing	42	40	Pickles or olives	41	34	Quickbread <sup>e</sup>	47	53
Grapefruit	40	32	Chocolate	41	38	Grits	40	46
Cucumber	40	39	Ketchup or tomato chili sauce	37	31	Bacon	40	40
Green pepper	40	40				Beans	36	38
Summer squash	38	37	Beef	37	40	Rice	32	35
Pear	37	30	Processed meat <sup>f</sup>	36	34	Watermelon	32	34
Fresh tomato	35	36	White potatoes	35	46	Liver	29	32
Fish (fresh or frozen)	35	33	Sweets/desserts <sup>g</sup>	34	42			
Brussels sprouts	35	38	Cheese sauce, white sauce, thick gravy	34	32			
Apples	35	32						
Oranges	34	37	Corn	30	36			
Winter squash	34	35	Pasta	27	28			
Apricots	33	24	Ice cream	24	34			
Bananas	32	33	Cheese and cheese dishes					
Vegetable soup	32	25		24	26			
Cottage cheese	31	25	Peanuts	22	22			
Strawberries	28	26						
Canned fish	26	28						
Cooked tomatoes	25	24						
Plums	24	24						
Shellfish	23	28						
Nuts (not peanuts)	22	21						
Sweet red pepper	22	22						
% Variance	5.2	4.5		3.3	2.9		2.3	3.4
Coefficient $\alpha$	0.77	0.74		0.64	0.63		0.62	0.65

<sup>a</sup>  $N = 1,866$  men.

<sup>b</sup>  $N = 1,913$  men.

<sup>c</sup> Includes potato chips, pretzels, crackers, and salted nuts.

<sup>d</sup> Includes roast pork, pork chops, fresh ham, and spare ribs.

<sup>e</sup> Includes sandwich or packaged luncheon meats, hot dogs, and meat spreads.

<sup>f</sup> Includes cakes, donuts, cookies, pies, and candy.

<sup>g</sup> Includes muffins, biscuits, and flour tortillas.

**Table 2** Sociodemographic and health behavior characteristics and intake of selected dietary factors for first and third dietary pattern tertiles in 3,544 adult male participants in the National Health Examination Follow-up Study, 1982–84

	Vegetable-fruit pattern tertiles		Red meat-starch pattern tertiles		Southern pattern tertiles	
	1	3	1	3	1	3
Mean (SD) age (y)	57 (15)	59 (14)	61 (14)	55 (14)	57 (14)	59 (15)
Black (%)	14	9	16	7	3	24
Region (%)						
Northeast	25	33	31	26	35	19
Midwest	28	24	19	33	36	14
South	26	15	24	16	7	45
West	21	28	26	25	22	22
Residence (%)						
Rural	44	34	33	44	33	48
Urban	37	38	42	32	38	36
Suburbs	19	28	25	24	30	16
Education (%)						
< high school	47	34	43	36	32	57
completed high school	32	29	26	35	34	24
>high school	21	37	31	29	34	20
Poverty-income ratio (%) <sup>a</sup>						
≤1	15	8	14	8	6	18
>1–3.5	60	53	53	63	57	60
>3.5	23	38	33	29	37	22
Current smoker (%)	36	23	27	35	28	34
Current multivitamin use (%)	20	27	26	21	25	20
“Much” exercise (%)	16	24	18	24	19	21
Median times per week						
Red meat	8.0	8.0	5.1	11.0	7.0	9.0
Dairy	9.0	11.0	8.0	13.0	10.2	9.5
Fruits and vegetables	13.9	38.2	21.5	27.0	20.1	29.0
Tomatoes	3.1	6.0	3.3	6.0	4.3	4.8
Pearson correlations						
Energy (kcal)		0.29		0.73		0.25
Total fat (g) <sup>b</sup>		−0.11		0.28		−0.02
Saturated fat (g)		−0.19		0.26		−0.09
Fiber (g)		0.63		−0.03		0.20
Vitamin A (IU)		0.47		−0.05		0.35
Vitamin C (mg)		0.56		−0.05		0.03
Calcium (mg)		0.14		−0.04		−0.07

<sup>a</sup> Ratio of family income to a Census Bureau-determined poverty threshold for household size and adult/child composition of family; a ratio <1 is considered to represent below poverty level. (*N* = 2,491).

<sup>b</sup> All nutrients were energy-adjusted using the residual method (36).

*Table 3* Adjusted relative risk (RR) estimates and 95% confidence intervals (CI) by dietary pattern intake for 3,616 adult male participants in the National Health Examination Follow-up Study followed 1982–84 to 1992

Dietary pattern	No. cases	Person-years	Minimal model <sup>a</sup>		Full model <sup>b</sup>	
			RR	95% CI	RR	95% CI
<b>Vegetable-fruit</b>						
Tertile 1	35	9,053	1.0		1.0	
Tertile 2	51	9,432	1.4	0.9–1.8	1.5	0.9–2.3
Tertile 3	45	9,359	1.2	0.7–1.8	1.2	0.7–2.0
P for trend <sup>c</sup>				0.72		0.64
<b>Red meat-starch</b>						
Tertile 1	61	8,823	1.0		1.0	
Tertile 2	38	9,431	0.8	0.5–1.2	0.7	0.5–1.2
Tertile 3	32	9,590	0.9	0.6–1.3	0.8	0.4–1.4
P for trend				0.47		0.37
<b>Southern</b>						
Tertile 1	45	9,562	1.0		1.0	
Tertile 2	43	9,606	0.9	0.6–1.3	0.9	0.6–1.4
Tertile 3	43	8,675	0.7	0.4–1.0	0.6	0.4–1.1
P for trend				0.06		0.08

<sup>a</sup> Models adjusted for age, race (white, black, other race), and design variables (dichotomized age (<65 vs. ≥65 years), poverty census enumeration district (residence vs. nonresidence), and family income (<\$3,000, \$3,000–\$6,999, \$7,000–\$9,999, \$10,000–\$14,999, and ≥\$15,000)).

<sup>b</sup> Models adjusted for above variables, and United States region (Northeast, Midwest, South, West), residence (rural, urban, suburban), education (<high school, high school completion, >high school), recreational sun exposure (little, occasional, frequent), recreational physical activity (little/none, moderate, much), smoking status (never, former, current), current alcohol intake (none, little, moderate, heavy), and energy intake (tertiles).

<sup>c</sup> P for trend was obtained for each pattern by including in the model a variable representing the median value for each tertile.