LaFontaine, Laura Jean, Dietary Quality in the United States: An Evaluation of the American Diet Using the Mediterranean Diet Index. Master of Public Health (Epidemiology), December 2003, 69 pp., 10 tables, 3 figures, reference list, 92 titles.

Mediterranean Diet Index (MDI) score is positively associated with increased longevity in diverse populations. High scores are characterized by high intake of fruit, vegetables, legumes, fish, and grains; low intake of meat and dairy; moderate alcohol intake; and high monounsaturated to saturated fat ratio. Dietary quality of 7,772 individuals ages 2 and older in the NHANES 1999-2000 sample was assessed using the MDI. Mean MDI scores were low overall and across all population groups. Highest mean scores (3.03–3.13) were observed in children 2-3 and adult men; lowest scores (2.76-2.77) were observed in adolescents. High scores were associated with foreign-born status, higher education level, increasing age in adults, and decreasing age in children.

DIETARY QUALITY IN THE UNITED STATES: AN EVALUATION OF THE AMERICAN DIET USING THE MEDITERRANEAN DIET INDEX

Laura Jean LaFontaine, B.S., M.B.A.

APPROVED:

Major Professor

Committee Member

Committee Member

Department Chair

Dean, School of Public Health

DIETARY QUALITY IN THE UNITED STATES: AN EVALUATION OF THE AMERICAN DIET USING THE MEDITERRANEAN DIET INDEX

THESIS

Presented to the School of Public Health

University of North Texas Health Science Center at Fort Worth

in Partial Fulfillment of the Requirements

for the Degree of

Master of Public Health

By

Laura Jean LaFontaine, B.S., M.B.A.

Fort Worth, Texas

December 2003

ACKNOWLEDGEMENTS

I would like to thank my committee members for their guidance throughout the thesis preparation process, particularly Sejong Bae, Ph.D., for much appreciated assistance with SAS programming. I would also like to thank Nancy Raper, Ph.D., and Sheryl Gerrior, Ph.D., of the United States Department of Agriculture for their assistance in providing access to the recipe database on which the NHANES 1999-2000 food codes are based. Finally, I would like to thank my dear husband, Kim, and my children, Claire and Peter, for their tremendous support and patience while I worked to complete this degree.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

"The underlying premise in contemporary nutritional epidemiology is that a person's long-term habitual diet has an impact on the occurrence of chronic disease."

Wilkens and Lee, 1998, p. 3099

Dietary differences between populations have been studied in an attempt to determine which aspects of human diets are responsible for higher rates of disease or greater longevity. Pivotal studies such as the comparison of cancer rates between Japanese living in Japan and Japanese migrants in the United States (Haenszel & Kurihara, 1968), Carroll's analysis in 1975 of per capita dietary fat consumption and age-adjusted breast cancer mortality in 39 countries, and the Seven Countries Study examining risk factors of coronary heart disease (CHD) (Keys, 1980) helped to determine the direction of future nutritional epidemiological investigations. While these studies demonstrated seemingly convincing links between diet and disease, identifying the specific aspects of diet responsible for heightened disease rates was not straightforward.

Initial investigations into diet/disease relationships focused on single nutrients or micronutrients and disease occurrence. Although these were the most common types of dietary analyses performed during the 1980s and 1990s, a minority of researchers attempted to investigate the impact of whole diets, or dietary patterns (Schwerin et al. 1981; Randall, Marshall, Graham & Brasure, 1990; Slattery, Boucher, Caan, Potter & Ma, 1998). They suggested that, because diets are complex combinations of numerous foods and nutrients, it was not possible to evaluate dietary effects by assessing the impact of isolated components. By the end of the 1990s and early 2000s dietary pattern analysis was identified as an important new direction in nutritional epidemiology (Willett, 2000; Hu, 2002).

Dietary pattern analysis may be performed '*a priori*' or '*a posteriori*'. *A priori* analysis assesses a set of dietary intake data against a pre-established, often published, dietary standard to determine the extent to which that dietary quality standard is being met in the population under study. Dietary guidelines published by the United States Department of Agriculture (USDA), the World Health Organization, and the American Heart Association have all been used in *a priori* dietary pattern analyses. In addition, comparison of dietary data to an index based on the parameters of a defined diet such as the Mediterranean Diet, is also considered *a priori* dietary pattern analysis.

In contrast, an *a posteriori* approach to dietary pattern analysis involves the use of multivariate statistical analysis techniques to identify dietary patterns inherent within a set of collected data. This is accomplished by searching for correlations either between foods consumed (correlations at the food item or food group level) or between the reported dietary habits of individuals (correlations at the level of the individual). Principal components (factor) analysis and cluster analysis are the two statistical methods most often used to identify dietary patterns *a posteriori*. Both techniques simplify expansive food intake data sets by revealing approximately 2-8 meaningful dietary patterns within the data. In contrast to the *a priori* methods that identify the extent to which dietary consumption in a population is nutritionally sound (i.e., identify dietary quality), factor and cluster analysis attempt to discover the eating patterns that occur most frequently among the sample population without regard to dietary quality or comprehensiveness.

Dietary patterns have been linked to socioeconomic and other demographic and lifestyle characteristics of populations in Europe (Haveman-Nies, Tucker, de Groot, Wilson & van Staveren, 2001; Whichelow & Prevost, 1996; Martikainen, Brunner & Marmot, 2003), Australia (Mishra, Ball, Arbuckle & Crawford, 2002), Asia (Woo et al. 2001), and North America (Beaudry, Galibois & Chaumette, 1998; Wirfält & Jeffery, 1997). Dietary patterns have also been found to be associated with coronary heart disease (Hu et al. 2000), colon and other cancers (Slattery et al. 1998, Siari et al. 2002), the metabolic syndrome (Williams et al. 2000), and osteoporosis (Tucker et al. 2002). In some instances both *a priori* and *a posteriori* analysis methods are applied to the same data set. Haveman-Nies et al. performed a combined analysis of dietary data from the United States Framingham Heart Study and the European

SENECA study using cluster analysis and two indexes, the Healthy Diet Indicator score based on 1990 World Health Organization guidelines, and a Mediterranean Diet Score. Both methods demonstrated associations with nutritional and lifestyle factors, and the authors concluded that the two diet analysis methods were complementary.

In this study, dietary data from a representative sample of the United States population, the 1999-2000 National Health and Nutrition Examination Survey (NHANES 1999-2000), will be analyzed using the Mediterranean Diet Index (MDI) (Trichopoulou, Costacou, Bamia & Trichopoulos, 2003) to assess the quality of the American diet. The health benefits of the Mediterranean diet were first postulated by Keys as a result of the Seven Countries research (Keys, 1980). Benefits of the Mediterranean diet were demonstrated in a prospective experimental study, the Lyon Heart Study, in which increased CHD survival was observed in participants who modified their diets in line with Mediterranean diet guidelines, compared to participants who followed a prudent Western diet (de Lorgeril et al. 1999). The MDI has demonstrated an association between dietary quality (evidenced by high MDI scores) and increased longevity in populations both in and away from the Mediterranean region (Trichopoulou et al. 1995, Lasheras, Fernandez & Patterson, 2000; Singh et al. 2002; Osler & Schroll, 1997; Kouris-Blazos et al. 1999). While the index is based on the traditional diet of the Mediterranean region, it is general in character, allowing it to be broadly applied in populations around the world, regardless of the specific foods that are typically eaten. The Mediterranean diet is characterized by relatively high intakes of fruit, vegetables, legumes, fish, and grains; relatively low intakes of meat and dairy products; a moderate intake of alcohol; and a relatively high ratio of monounsaturated to saturated fat. These nine dietary qualities form the basis of the index.

The quality of the American diet, as represented by the NHANES 1999-2000 data set, has been previously measured against an external standard, the Healthy Eating Index (HEI), which is based on USDA Food Consumption Guidelines as represented by the USDA Food Guide Pyramid (Basiotis, Carlson, Gerrior, Juan & Lino, 2002). However, analysis of the same United States dietary intake data against the MDI, a second, quite different external standard, may provide further useful insight into the current quality of the American diet. In this study, MDI scores of all persons age 2 and older will be calculated and analyzed against the demographic variables age, gender, race, education level, and country of birth, to assess the quality of the American diet overall, and within demographic subsets of the sample population.

CHAPTER II

LITERATURE REVIEW

"Diet is a complex exposure variable, which calls for multiple approaches to examine the relationship between diet and disease risk."

Hu 2002, p. 8

Epidemiology of Chronic Diseases

Mortality due to chronic diseases, principally heart disease, cancer, stroke, and diabetes, increased tremendously during the 20th century concurrent with the decline in mortality due to infectious diseases. Among the ten leading causes of death in the United States in 1900, five were chronic diseases causing 29.4% of deaths; by 1995, six were chronic diseases causing 70.2% of deaths (McKenna, Taylor, Marks & Koplan, 1998).

Chronic diseases are "generally characterized by uncertain etiology, multiple risk factors, a long latency period, a prolonged course of illness, non-contagious origin, functional impairment or disability, and incurability" (McKenna et al. 1998, p.3). While the causes of chronic diseases are unclear, risk factors identified to date include many environmental and lifestyle factors, including cigarette smoking, physical inactivity level, exposure to pollutants, and obesity. Genetic and physiological factors also increase the risk of chronic diseases (McKenna et al.). In addition, diet has been identified as a factor in chronic disease occurrence (Kushi & Foerster, 1998).

Goals in chronic disease control include reducing prevalence through prevention, delaying onset of disability, alleviating disease severity, and prolonging life (McKenna, et al. 1998). One model of the future chronic disease burden speculates that with a growing ability to extend lives, chronic diseases will occupy more time in a person's life, while another model predicts more disease free years due to healthier lifestyles and a subsequent delay in onset of chronic diseases (McKenna, et al.). An emerging model of chronic disease epidemiology, the life-course approach, attributes chronic disease etiology to the long-term effects of factors from gestation through adulthood (Kuh & Ben-Shlomo, 1997). In this model, maternal diet is an important factor in the health of the developing fetus (Perry, 1997), and diet of the family is emphasized as being critical in the formation of healthy behaviors such as the establishment of dietary preferences and patterns (Kuh, Power, Blane & Bartley, 1997). Studies of diet in relation to chronic disease occurrence and mortality have demonstrated the importance of dietary quality throughout adulthood (Keys, 1980; Osler & Schroll, 1997; Kant, Schatzkin, Graubard & Schairer, 2000).

Nutritional Epidemiology and Dietary Links to Chronic Disease

Early findings on diet/disease relationships resulted from ecologic and migrant studies, as well as key cohort studies. In an ecologic examination of data from 39 countries, a correlation between age-adjusted breast cancer mortality and per capita dietary fat intake was discovered, with lowest mortality and lowest total fat intake in Thailand, and highest mortality and highest total fat intake in the Netherlands, United Kingdom, and Denmark (Carroll, 1975). Haenszel and Kurihara (1968) compared cancer mortality rates in original (native Japanese) and migrant (Japanese immigrants to the United States) populations, and found that cancer incidence rates were associated with the environment in which one lived (one's new location) rather than one's genetic heritage. In the Seven Countries Study, Keys et al. (1981) explored connections between diet, plasma cholesterol, arterial blood pressure, and other risk factors, and all-cause and CHD mortality in cohorts of men from 16 regions in 7 countries (United States, Yugoslavia, Japan, Finland, Italy, the Netherlands, and Greece). The countries represented were chosen, in part, in an effort to ensure dietary diversity between cohorts (Keys, 1980). This study provided a first comparative view of the Western, Japanese, and Greek diets in connection with mortality, and identified saturated fat intake as a key risk factor in early death (Keys, 1980).

Following these early discoveries, the majority of new research into diet/disease relationships investigated the potential effects of single nutrients and micronutrients. For example, numerous studies followed up on Carroll's ecologic study results and investigated the relationship between total fat intake and breast cancer (Goodwyn & Boyd, 1987). While results from case-control studies were more promising, cohort studies provided mixed results (Goodwyn & Boyd; Hunter, 1999). In their review of studies on dietary fat and breast cancer, Goodwyn and Boyd (1987) found no evidence of a causal link. Hunter (1996) also found no evidence for a relationship between total dietary fat intake and breast cancer following a pooled analysis, and relative risk estimates for intake of saturated, monounsaturated, and polyunsaturated fats were also near unity. In a paper in support of a possible causal role for dietary fat in the occurrence of breast cancer, Greenwald (1999) concluded that the evidence for a causal link was inconsistent, but that lack of evidence may be due to measurement error and limitations in study designs. Weisburger (1997) analyzed the mechanisms behind type of fat and promotion of tumor development, and concluded that saturated fats, monounsaturated oils, and n-3 and n-6 polyunsaturated oils all produce different effects in the body. His recommendation was to increase monounsaturated and n-3 polyunsaturated oils, while decreasing n-6 polyunsaturated oils, in addition to adding other protective foods including bran, fruits and vegetables, and tea. This recommendation to adjust the ratio of oil intake by type, along with adding other protective foods to the diet, represented a change in the approach toward analysis of diet/disease relationships. The movement away from analyzing a single dietary component to looking at more global diet-level effects was not unique to investigations into the etiology of breast cancer.

Research into the etiology of osteoporosis investigated the effect of calcium intake with mixed results (Reid, Ames, Evans, Gamble & Sharp, 1993; Feskanich, Willett, Stampfer & Colditz, 1997; Cumming et al. 1997; Cumming & Nevitt 1997). More recently, bone mineral density (BMD) and dietary patterns were assessed in the Framinghan Osteoporosis Study (Tucker et al. 2002) in which the authors found significantly greater BMD in men eating the dietary pattern high in fruits, vegetables, and cereal ($p=0.05$) compared to men in other dietary groups. Women in the candy pattern had significantly lower BMD at the radius than all but one other group ($p<0.01$). The authors offer the following rationale for dietary pattern analysis in studying the etiology of osteoporosis:

> Bone is a complex living tissue, and it is probable that a wide spectrum of micronutrients contributes to its maintenance. Because of that complexity, examining nutrients individually may sometimes be misleading.

> Nutrients tend to be packaged together in foods, and therefore associations seen with a single nutrient may, in fact, be caused by a more complex constellation of other nutrients consumed contemporaneously. Conversely, adjusting for other nutrient contributors to bone mineral density may make it difficult to see a true association because of their close association with one another in a healthy diet. (Tucker et al. 2002, p. 245)

The association between aspects of diet and CHD were established early on, in large part because of studies like the Seven Countries Study (Keys, 1980) and the Western Electric Study (Shekelle et al. 1981). Kushi and Kottke (1990, p. 396) summarized the evidence for a causal relationship between dietary factors and CHD and were able to make the following dietary recommendations:

1. Adjust energy intake to maintain desirable weight;
2. Reduce total fat intake to 30% of calories or less;
3. Reduce saturated fats to 10% of calories or less;
4. Increase polyunsaturated fats to no more than 10% of calories;
5. Reduce cholesterol to 300 mg or less per day.

The second, third and fifth recommendations in this list are all included in the current USDA food consumption guidelines (USDA, 2000). Hu et al. (1997) analyzed data from the Nurses' Health Study cohort for associations between dietary fat and the risk of CHD in women and concluded that it would be more beneficial to replace saturated and trans unsaturated fats with unhydogenated monounsaturated and polyunsaturated fats than to reduce overall fat intake. In 2001, Fung, Willett, et al. reanalyzed the same data set (Nurses' Health Study cohort), this time investigating the presence of associations between dietary patterns (rather than dietary fats) and CHD risk. Factor analysis yielded 2 principal dietary patterns, one characterized by higher intake of fruit, vegetables, legumes, fish, poultry and whole grains, the other characterized by higher intakes of red and processed meats, sweets and desserts, french fries, and refined grains. After adjustment for CHD risk factors, a protective effect was found to be associated with the highest quintile compared to the lowest quintile of the first dietary pattern (RR 0.76, 95% CI 0.60-0.98), while a 46% increased risk of CHD was found to be associated with the highest quintile compared to the lowest quintile of the second dietary pattern (RR 1.46, 95% CI 1.07-1.99). The authors concluded that "a diet high in fruits, vegetables, whole grains, legumes, poultry, and fish, and low in refined grains, potatoes, and red and processed meats may lower risk of CHD" (Fung, Willett, et al., p. 1857). In the case of CHD, therefore, both isolated aspects of diet (dietary fat) and whole dietary patterns have been shown to be associated with disease occurrence.

The earliest account of dietary pattern analysis, a comparison of nutrient intake to an index for the purposes of measuring pregnancy outcomes, was performed by Burke et al. in 1943 (Kant, 1996). Since then, many additional index comparisons have been performed, some based on nutrient intake, others on food or food-group intake (Kant). These comparisons have been made for the purpose of assessing nutritional adequacy, socio-demographic and lifestyle factors, and all-cause and cause-specific mortality in relation to dietary quality.

Schwerin et al. (1981) was the first to use factor analysis to derive dietary patterns from food intake data (Ten-State Nutrition Survey 1968-1970 and Health and Nutrition Examination Survey I 1971-1974), and found associations between dietary patterns and presence of clinical symptoms. Akin, Guilkey, Popkin & Fanelli (1986) initiated the use of cluster analysis, aggregating data on the dietary habits of individuals (Nationwide Food Consumption Survey 1977-1978), to identify dietary patterns and assess the nutritional adequacy and socio-demographic factors associated with each pattern. In subsequent years, factor analysis and cluster analysis were performed by nutrition researchers in the United States and elsewhere, often, but not always, demonstrating positive associations between dietary patterns and health outcomes (Table 1, *A Posteriori* Analyses of Dietary Patterns).

In addition, while a number of large randomized controlled trials of individual nutrients, including fiber, beta carotene, and antioxidant vitamins, have been unsuccessful in showing a beneficial effect (Jacques & Tucker, 2001), clinical investigations of dietary patterns (DASH trial on hypertension - Appel et al. 1997 and the Lyon Heart Study on CHD - de Lorgeril et al. 1999), have demonstrated significant improvement in health outcomes to be associated with specific diets. In their review of the use of dietary pattern analysis, Jacques and Tucker (p. 2) concluded that "The use of dietary patterns should ultimately prove to be an informative and powerful means to augment our understanding of the role of diet in chronic disease."

Identifying Dietary Patterns

A posteriori identification of dietary patterns is accomplished by application of multivariate statistical methods to a set of dietary data in an attempt to identify correlational patterns within the data set. Principal components analysis (PCA), a type of factor analysis, and cluster analysis are the two multivariate methods most commonly used for this purpose (Hu, 2002). Multiple correspondence analysis (MCA), a technique developed and applied primarily in France, has some features that may make it especially suitable for analysis of dietary data (Guinot et al. 2001). The three methods are similar in that they all aim to simplify a data set by means of data aggregation so as to facilitate its interpretation. In traditional analyses of a single dietary factor, controlling for nutrients or foods that might confound or interact with the variables of interest is necessary, but very difficult because of the colinearity of dietary data. In contrast, PCA, MCA, and cluster analysis take advantage of, rather than attempt to control for, the natural multicolinearity of dietary data. The aggregated data groupings (factors, components, clusters) which are generated by these analyses are the dietary patterns. In general 2 to 8 dietary patterns are identified in a data set, regardless of the multivariate method used. The differences, strengths, and limitations of each analysis method are described in the sections that follow, along with specific examples of their application in analyzing dietary information.

i. Principal Components (Factor) Analysis

Analysis of dietary patterns was first performed by Schwerin et al. (1981) when PCA was used to identify the dietary patterns of 32,086 participants in two large, national surveys, the Ten-State Nutritional Survey 1968-1970 (Ten-State) and the Health and Nutrition Examination Survey 1971-1974 (HANES I). Both surveys collected food intake data by means of 24-hour recalls, resulting in identification of over 1,000 different food items. Schwerin et al. reduced the food items, in a step-wise fashion, into 15 food categories (Table 2, Comparison of Analytical Characteristics of Principal

Components Analysis in Three Early Applications) on which the factor analysis was performed. The solution was rotated to improve interpretability of the analysis; an orthogonal (Varimax) rotation was selected. Factors loading high (greater than 0.35) either positively (representative of very frequent consumption of that food group in the pattern) or negatively (representative of very infrequent consumption of a food group in the pattern) were retained. A standard eigenvalue cut-off point of 1.0 was used, and seven meaningful dietary patterns were identified out of the factor solution, explaining 55.3% of the variability in the data set. Using multiple discriminant analysis, Schwerin et al. confirmed that the seven eating patterns were significantly different from one another ($p<0.001$). Patterns (factors) are distinguished from one another by a distinct set of factor weights. Each individual's food intake data were scored against the factor weight to determine the factor or dietary pattern that best matched their habitual food consumption (i.e., the pattern for which an individual achieved the highest factor score, was the pattern to which that person was assigned). Schwerin et al. then correlated each dietary pattern with disease symptoms, as reported in the survey data. The presence of disease symptoms was found to be most associated with the "more nonsugary beverages, less dairy" pattern ($p<0.05$), and least associated with the "more dairy and soup, less sugary foods and beverages" pattern ($p<0.05$).

In 1990, Randall et al. applied PCA to food-frequency questionnaire data collected from 1,475 male and 780 female controls in a case-control study, the Western New York Diet Study 1975-1986 (WNYDS). The food-frequency questionnaire contained 110 items, and Randall et al. chose not to collapse the data into a smaller set of food categories, but rather ran the analysis on the full 110 food items. The analysis was performed separately by gender and identified 9 patterns in each group, explaining 22% of the variability in men and 23.1% in women. Like Schwerin et al., Randall et al. rotated the solution orthogonally for purposes of interpretation, however they used slightly different factor loading (<-0.2 and >0.3) and eigenvalue (1.5) cut-off points, to determine inclusion of food items within the dietary patterns, and inclusion of dietary patterns within the solution, respectively. The authors analyzed the dietary patterns for the presence or absence of nutrients believed important in cancer prevention, as well as for associations with demographic characteristics. The patterns somewhat reflected the ethnic heritage of the study participants as evidenced by the pattern names (Table 2). The authors found that nutrients of interest in cancer prevention, such as energy, fat, fiber, and vitamins A and C, were not correlated with a single dietary pattern. Rather, patterns tended to include both cancer risk factors and protective items, leading them to conclude that assessment of cancer risk on the basis of single nutrients could likely be confounded by differences in underlying dietary patterns.

Slattery et al. (1998) performed PCA of the dietary data (food-frequency questionnaires) of 2,389 male and 2,014 female participants in a population-based case-control study on colon cancer conducted in Northern California, Utah, and Minnesota. Individual food items were consolidated into 35 food groups, and the factor analysis was run separately for men and women. The solution generated 6 patterns for each gender, explaining 36.9% and 34.3% of variation for men and women, respectively. Varimax rotation was performed; eigenvalue and factor loading cut-off points were different from, but similar to, those used in the analyses performed by Schwerin et al. (1981) and Randall et al. (1990) (Table 2). Slattery et al. identified two primary patterns, loading first and second in the solution and therefore explaining the greatest proportion of the variability, which they named the 'Western diet' and the 'prudent diet'. The Western diet included high levels of red meat, processed meat, fast food, refined grains, and sugar-containing foods, and low levels of vegetables (other than potatoes) and fruits, with canned fruit being predominant. The prudent diet included consumption of all types of fruits and vegetables, fish and poultry were consumed more frequently than red or processed meats, and low amounts of high-sugar foods were eaten. The Western diet was found to be associated with higher body mass index (BMI), higher total energy intake, higher cholesterol, and increased risk of colon cancer (OR for men 1.96, 95% CI 1.22-3.15; OR for women 2.02, 95% CI 1.21-3.36). The prudent diet was associated with higher activity level, lower BMI, and was protective (OR for men 0.63, 95% CI 0.43-0.92; OR for women 0.58, 95% CI 0.38-0.87). The associations that were strongest, and that are reported here for both dietary patterns, were among people diagnosed prior to age 67 years.

Principal components analyses of dietary data were carried out by others in North America (Nicklas, Webber, Thompson & Berenson, 1989; Beaudry et al. 1998) and Europe (Whichelow & Prevost,

1996) in the same timeframe. In some applications of PCA, dietary patterns identified are less intuitively representative of common, complete diets, because they are based on only a few food items (i.e., Schwerin et al. (1981) "soup and dairy" pattern, or Randall et al. (1990) "fruit" pattern). The Slattery et al. (1998) Western and prudent dietary patterns include a broader range of foods and as a result more closely resemble complete diets. Between 1998 and 2003 numerous data sets were analyzed for the presence of dietary patterns (Table 1). Associations were found between dietary patterns and CHD in men (Hu et al. 2000; Fung, Rimm, et al., 2001) and women (Fung, Willett, et al. 2001); and between dietary patterns and features of the metabolic syndrome (Williams et al. 2000), a variety of cancers (Siari et al. 2002), mortality (Osler, Heitmann, Gerdes, Jørgensen & Schroll, 2001), and colon cancer (Fung et al. 2003). However in other analyses of dietary patterns and colorectal cancer (Terry, Hu, Hanson & Wolk, 2001), breast cancer (Terry, Suzuki, Hu & Wolk, 2001), and CHD (Osler et al. 2002) associations were either of borderline significance or were not found.

In PCA, the correlational patterns identified are based on individual food items or, more often, food groups. A strength of PCA and of all of the *a posteriori* methods, is their ability to synthesize complex food intake data into a few key patterns, providing a more comprehensive way of analyzing interactions between diet and health outcomes. A criticism of PCA is the series of arbitrary decisions that must be made by the researcher, causing the analysis to be difficult, if not impossible, to reproduce (Martinez, Marshall & Sechrest, 1998). These decisions include 1) Should the food item data be collapsed into broader food groups, and if so, how many groups should there be and how should they be configured? 2) How many of the factors in the factor solution are meaningful and should be retained? 3) Should the factor solution be rotated for improved interpretability, and if so, in which direction? and 4) Once the factors to be kept are chosen, how should they be labeled? Each of the decisions identified by Martinez et al. is reviewed below. The implications of these decisions are highlighted, and the degree to which a standard approach appears to have developed in nutritional epidemiologic usage, as evidenced by this literature review, is discussed.

First, collapsing the data into food groups makes the data more manageable and can focus the analysis on food groups or food items of particular interest in the analysis. With the exception of Randall et al. (1990) who conducted the factor analysis on all 110 food items, all researchers included in this review chose to collapse the food items into a smaller number of food groups, unless the food-frequency questionnaire collected data on only a limited number of foods at the outset (Williams et al. 2000). Among the studies reviewed, the greatest number of food groups analyzed using PCA was 110 (Randall et al.), while the smallest number was 15 (Schwerin et al. 1981). In general, food items from food-frequency questionnaires are collapsed into groups when appropriate, but individual foods (eggs, pizza, soup) or foods that might represent a pattern (garlic, beer, french fries) are included in the analysis individually (Hu et al. 2000). In the studies reviewed, the majority of PCA were based on 20 to 40 food items. The number of groups or items on which the factor analysis is performed affects the level of variability explained by the solution. In general, if fewer items are included in the analysis, a greater proportion of the variability in the data set is explained by the factor solution (Hu et al. 2000, Osler et al. 2001).

Second, the choice as to the number of factors (dietary patterns) to retain in the factor solution is another key decision. Factors are accepted into the solution based on the degree to which they are able to describe a distinct and meaningful dietary pattern. Generally an eigenvalue cut-off point is defined based upon a review of the solution, and all factors with eigenvalues below that point are deemed unimportant in explaining meaningful dietary differences and are discarded from the analysis. While an eigenvalue cut-off point of 1.0 is standard (Armitage, Berry and Matthews, 2002), Randall et al. (1990) and Slattery et al. (1998) chose to use slightly higher eigenvalue cut-points of 1.5 and 1.25, respectively. Maskarinec, Novotny and Tasaki (2000) and Williams et al. (2000) used scree plots to assist in determining which factors or dietary patterns were contributing most to the solution and ought to be kept. In a scree plot, the point at which the slope changes dramatically (there is a defined elbow in the plot) is the point at which factors contribute so little to the solution as to be not worth retaining (Williams et al.). Hu et al. (2000) used eigenvalues, scree plots, and interpretability of the factors to determine which factors to retain, and the use of one or all of these indicators to assist in determining the number of factors in the solution was universally reported by all reviewed authors.

Third, in the studies reviewed (Table 1), all researchers chose to rotate the factor solution orthogonally, causing rotation to no longer be an important point of analytical differentiation. It is possible that orthogonal rotation has become a de facto standard in principal components analysis of dietary patterns since it is consistent with the desire to achieve factors (patterns) that are uncorrelated with one another.

Fourth, the labeling of dietary patterns projects the researcher's understanding of what in the pattern is "the latent variable that might have produced the observed factor loadings" (Martinez et al. 1998, p. 18). Martinez et al. point out that while some of the Slattery et al.(1998) patterns appear to be well-named, like 'Western' and 'prudent' diet, others do not, for example, coffee loads higher on the 'drinker' pattern than it does on the 'coffee and roll' pattern, but seems to be accorded a more important role in the latter because of the pattern name. Nevertheless, and perhaps as a result of the observations of Martinez et al. in 1998, some consistency in the labeling of dietary patterns has occurred. After Slattery et al. identified the Western and prudent patterns, variations of these same patterns were found by later researchers, both in Europe and in the United States, and, in some cases, given the same labels (Hu et al. 1999 and 2000; Fung, Rimm, et al. 2001; Fung, Willett, et al. 2001: Osler et al. 2001). This consistency in dietary pattern identification and labeling allowed for greater ability to compare findings across studies.

Martinez et al. (1998) also question whether the results in a factor analysis solution are reproducible or might lead to inconsistent results in nutritional research, and recommend that confirmatory factor analysis be used as a means of testing the robustness of the factor solution. Confirmatory factor analysis was performed by Maskarinec et al. (2000) by rerunning the analysis on random halves of the data set to confirm that the same solution was generated each time. In addition, discriminant analysis has been used to test whether the dietary patterns identified are in fact, unique, uncorrelated patterns (Schwerin et al. 1981). Hu et al. (1999) demonstrated the reproducibility and validity of dietary patterns defined by factor analysis by assessing the presence of patterns first in a large population sample (Health Professionals Follow-up Study, n=51,519) based on data collected by means of two food-frequency questionnaires administered one year apart. These results were compared to dietary pattern assessments of two, 1-week diet records administered 6-7 months apart from a small sub-sample of the same population (n=127). Two primary dietary patterns, Western and prudent, were identified in each analysis. Correlations between the food-frequency questionnaire data were 0.70 and 0.67 for the prudent and Western diet patterns, respectively. Correlations between the 1-week diet record and the food-frequency analyses ranged from 0.45 to 0.74, and correlations between factor score and biomarkers were in the expected direction (Hu et al). The cautions raised by Martinez et al. in 1998 perhaps served the research community by guiding it to a rigorous and fairly uniform application of PCA in subsequent years, helping to establish it as a viable method of nutritional data analysis.

ii. Multiple Correspondence Analysis

Multiple correspondence analysis (MCA) is a type of principal components analysis performed with categorical (dummy) variables (Armitage et al. 2002). A method originated in France, its use has been limited outside of that country, however MCA may have some advantages over traditional principal components analysis because 1) its categorical nature makes it appropriate for use in positively skewed nutritional data and 2) it results in useful graphical displays that assist in understanding the relationships between subjects, dietary patterns, and disease (Guinot et al. 2002).

In an analysis of the SU.VI.MAX. cohort nutrition data, Guinot et al. (2002) identified a high degree of positive skewness. After determining that linear correlation analysis (i.e. principal components analysis) methods were not appropriate for this data set, the authors chose to dichotomize the nutrition data "in order to increase the strength of the links between the food items" (Guinot et al.). It was hypothesized that links would exist between food items that were over-consumed, so the median was selected as the cut-off point. Consumption over the median was scored '1', and zero consumption or under-consumption was scored '0'. MCA was performed separately by gender using the binary data. Three principal factors 'luxurious', 'dietetic', and 'bread and butter' were retained in the solution for each gender, explaining 78% and 79% for men and women, respectively. The 3-dimensional relationships between the three dietary patterns were graphed in two, 2-dimensional displays for each gender, and showed the foods contributing to each pattern and their relative position to other foods both in the pattern of interest and in 'adjacent' patterns.

Because nutritional data are frequently skewed (Milton, 1998) and because PCA requires that data be normally distributed, the binary quality of MCA may be particularly well-suited to dietary pattern analysis. Maskarinec et al. (2000) performed PCA on their food-frequency data investigating diet and obesity, but reported non-normal distributions in a number of food item variables. They were able to correct for this in some food items by logarithmic transformation, but in other foods the skewness was not correctable, requiring these variables to be discarded and reducing the scope of their analysis. Other reviewed studies that used PCA did not mention skewness of the data, so the degree to which it is an issue in PCA is not known.

MCA results generate coordinate values that summarize the relationships between subjects and food categories, allowing these associations to be displayed graphically (Guinot et al. 2002). These displays can be helpful in understanding the relationships between foods within a dietary pattern, as well as between the dietary patterns themselves. Unfortunately, MCA has not been widely applied, and experience with this method of analysis is limited.

iii. Cluster Analysis

A key feature distinguishing cluster analysis from factor analysis is the manner in which the data are aggregated. Unlike factor analysis in which the unit of analysis is the individual food group or food item, in cluster analysis the unit of analysis is the individual (Akin et al. 1986). Individuals are aggregated into clusters based on the similarity (correlation) of their consumption patterns. As a result of this distinction, one of the advantages of cluster analysis is that the dietary patterns identified are actual patterns present in the population. If the analysis is performed in a representative sample of the population, prevalence rates of dietary patterns in that population may be calculated (Pryer et al. 2001). In addition, nutrient profiles may be calculated for each pattern and used in educating on malnutrition risk or other diet related health concerns (Akin et al.).

Akin et al. (1986) used cluster analysis to identify dietary patterns among the elderly participants of the Nationwide Food Consumption Survey (NFCS 1977-1978). More than 2,800 food items were classified into 20 categories that in particular differentiated food items on the basis of the fat and cholesterol content, and vitamin A and C content. Akin et al. used cluster analysis to identify eight dietary patterns for each gender, and calculated the nutrient profile of each pattern. Ethnicity and geographic area of residence were the socioeconomic factors most associated with dietary pattern differences.

Dietary intake of Seven Countries Study participants from two rural, Italian villages was analyzed using cluster analysis and correlated with twenty-year mortality data (Farchi et al. 1989). The authors state that "clusters were obtained in which differences in diet within groups were much smaller than between groups and impressive differences in mortality rates were found between groups" (Farchi et al., p. 1101). In this analysis, the categories analyzed were not food items, but rather nutrient categories (protein, carbohydrates, and 3 types of fatty acids) and alcohol. After 20 years of follow-up, the pattern unique for its high polyunsaturated fat content was associated with lowest CHD, stroke, and all cancers mortality, while the pattern lowest in all types of fat, highest in carbohydrates, and high in protein was associated with the lowest mortality due to stomach and lung cancer.

Millen et al. (1996) analyzed the dietary patterns of participants in the Framingham Offspring/Spouse Study using cluster analysis and then compared the patterns generated with USDA Food Guide Pyramid recommendations. Distinct dietary patterns were generated using cluster analysis that varied from national recommendations. The authors conclude that dietary pattern analysis can "provide insights for developing behavioral interventions to improve food and nutrient intake" (Millen et al.).

Quatromoni et al. (2002) demonstrated the reproducibility of the cluster analysis technique using discriminant analysis. Another method of confirming the robustness of a cluster analysis solution is, as with PCA, to rerun the analysis on random halves of the data to confirm that the same solution is generated each time (Martikainen et al. 2003). Millen et al. (2001) validated cluster analysis methodology for dietary pattern analysis by comparing analysis of food-frequency questionnaire data (Framingham Offspring/Spouse Study) with analysis of independent 3-day food record data, as well as with biological and behavioral risk factors. Cluster analysis successfully identified non-overlapping groups of participants with distinct dietary patterns in the two sets of food intake data, that further were shown to be associated with different levels of nutrient intake and disease risk factors (Millen et al.).

Measuring Dietary Quality

In contrast to the multivariate statistical techniques described above, comparison of food intake data to an index is an *a priori* method of dietary pattern analysis because the pattern of interest (represented by the parameters of the index) is pre-established and based on an external standard of nutritional quality, rather than derived from the data itself. *A priori* analysis by comparison to an index such as the HEI based on USDA Food Consumption Guidelines (McCullough, Feskanich, Rimm, et al. 2000; McCullough, Feskanich, Stampfer, et al. 2000), or the MDI based on the traditional diet of Greece (Trichopoulou et al. 1995) can be used to determine what proportion of the data sample is compliant with the dietary parameters represented by the index. If health outcome data are part of the data set, the analysis will also show whether or not there are health implications associated with compliance (or lack of compliance) to the diet or dietary guidelines.

i. Scores based on published dietary guidelines

Dietary guidelines are published by government agencies and health promotion organizations such as the USDA, the National Academy of Sciences, and the World Health Organization, as well as disease prevention organizations such as the American Heart Association. Guidelines are based on current understanding (at the time of publication) of what constitutes a healthy diet, or what dietary practices might help reduce the risk of a particular disease, such as heart disease or hypertension. Guidelines are often a series of recommendations that address both the quantity and variety of food and nutrient intake.

Indexes have been created on the basis of food consumption recommendations as a way of measuring or scoring compliance with these recommendations in a population. The HEI measures adherence to the USDA recommendations portrayed by the food guide pyramid, and other indexes have been created to measure other specific recommendations (Table 3, Comparison of Indexes of Dietary Quality). Kant, Schatzkin, Harris, Ziegler and Block (1993) measured dietary quality based on the recommendation to "eat a variety of foods" by scoring diets on the basis of the number of food groups consumed on a daily basis, and found that decreased diet diversity was associated with greater all-cause mortality, and cause-specific (CVD and cancer) mortality (Kant, Schatzkin & Ziegler, 1995). Patterson, Haines & Popkin, (1994) developed the Dietary Quality Index (Table 3) based on a broader set of dietary recommendations and used this index to analyze Nationwide Food Consumption Survey data. The authors found that important measures of dietary quality such as fiber and vitamin C intake were associated with high index scores, while such associations were not found with single dietary factors such as total fat intake. As a result, Patterson et al. emphasized the importance of multidimensional tools to measure dietary behavior. Later, Huijbregts et al. (1997) used the Healthy Diet Indicator (Table 3), an index similar to that used by Patterson et al., to assess 20-year mortality in men in Finland, Italy, and the Netherlands and found the relative risk for the healthiest diet group compared to the least healthy diet group to be 0.87 (95% CI 0.77-0.98).

In an effort to measure the impact of the most recommended foods on longevity, Kant, Schatzkin, Graubard and Schairer (2000) developed the Recommended Food Score (RFS) and applied it to weekly dietary intake of participants in the Breast Cancer Detection Demonstration Project. The RFS grants one point for each recommended food category present in one's diet over the course of a week, so that a higher score indicates consumption of a greater number of different, recommended foods. The list of 23 recommended food categories includes, 6 fruit and fruit juice categories, 1 legume category, 8 vegetable categories, 1 poultry category, 1 fish category, 4 grain categories, and 2 low fat milk categories (Table 3). After median follow-up of 5.6 years, women with greater consumption of foods from the recommended foods list had lower relative risk of all-cause mortality. After adjustment for education, ethnicity, age, BMI, smoking status, alcohol use, physical activity, hormone replacement therapy, and history of disease, the relative risk compared to those with a median score of 7.0 (seven different recommended foods consumed per week) was as follows: median score of 10.0, RR 0.82 (95% CI 0.73-0.92); median score of 12.0, RR 0.71 (95% CI 0.62-0.81); and median score of 15.0, RR 0.69 (95% CI 0.61-0.78); p for trend <0.001.

In a manner similar to Kant et al. (2000), McCullough, Feskanich, Rimm, et al. (2000) and McCullough, Feskanich, Stampfer, et al. (2000) analyzed the extent to which the diets of members of the Nurses' Health Study (women) and the Health Professionals Follow-up Study (men) met USDA dietary intake guidelines by assessing food intake against the HEI. This analysis showed that not only did the majority of participants NOT meet the daily recommended dietary requirements, but those that did, did not benefit from reduced (heart) disease risk. Study results led the researchers to question the value of the USDA dietary guidelines as currently articulated.

Limitations of the index method of dietary pattern analysis are that the index itself is limited by the level of current knowledge regarding diet/disease relationships. There are also subjective decisions made with regard to defining cut-off points and determining the individual score components which give more or less weight to certain dietary components. These decisions are biased by current evidence regarding dietary impacts on health. However, once the dietary factors of a disease have been identified, index comparisons might be an ideal way to evaluate the dietary guidelines themselves to determine if they may be adequate to prevent or reverse the course of a disease (Hu, 2002), or to maintain the health of the general population (McCullough, Feskanich, Rimm, et al. 2000; McCullough, Feskanich, Stampfer, et al. 2000).

ii. Scores and intervention studies based on established diets

Nube, Kok, Vandenbrouke, van der Heide-Wessel & van der Heide, (1987) developed an index based on a 'prudent' diet which awarded one point for consumption levels in ten categories. The score was based on daily, weekly, or monthly servings and rewarded regular consumption of brown bread, fruit, vegetables, fish, porridge or yogurt, and potatoes; limited consumption of milk, meat, and eggs; and avoidance of white bread. The authors found that after adjusting for confounders, higher diet scores in men were associated with significantly greater 25-year survival (survival rates were 50.4% [se 2.0%] versus 43.2% [se 2.0%] in the high and low dietary score groups, respectively; $p=0.01$). However, no difference in survival rates was seen among women across dietary score groups (survival rates were 68.7%

se=1.9%, 68.9% se=1.7%, and 66.1% se=2.4%, for low, moderate, and high dietary score groups, respectively, trend not significant, p=0.23).

The MDI is perhaps the most widely applied index based on an established diet. It was first developed by Trichopoulou et al. (1995) and is based on the traditional diet of Greece. It favors consumption of plant foods (vegetables, fruits, nuts, legumes, and cereals) over animal foods (meat and dairy), as well as a high monounsaturated to saturated fat ratio, and moderate alcohol consumption. In its original application, Trichopoulou et al. found high index scores to be associated with survival in an elderly, Greek population. The rate ratio associated with a one unit increase in diet score was 0.83 (95% CI, 0.69-0.99). Individual food groups that contributed to the scale score, were not generally associated with risk with the exception of dairy intake, which was associated with a 4% increase in risk of death for every 20 gram increase in daily consumption, after adjusting for energy intake.

Osler and Schroll (1997) also measured mortality in an elderly population in Denmark against a slightly modified MDI in which starchy root vegetables (potatoes) were grouped with vegetables, rather than with cereals, and legumes were also grouped with vegetables. Overall index scores were found to be positively associated with reduced risk of death (Osler & Schroll). The rate ratio associated with a one unit increase in diet score was 0.79 (95% CI 0.64-0.98). Similar results were demonstrated in an elderly, institutionalized, Spanish population (Lasheras et al. 2000) and in an ethnically mixed population in Australia (Kouris-Blazos et al. 1998). Woo et al. (2000) found the Mediterranean dietary pattern as measured by the MDI to be prevalent among Chinese populations living in San Francisco, Sydney, Hong Kong, and Pan Yu (rural China); score differences were noted by geographic region, age and gender, with highest scores measured in Pan Yu for both men and women. The high scores in Pan Yu were attributed in part to the area's rural location and relatively lower income levels, which may have precluded high protein (meat) intake.

In 2003, Tichopoulou et al. revised the MDI slightly by adding fish as its own independent food group and applied the index to the Greek arm of the European Prospective Investigation into Cancer and Nutrition (EPIC) study (n=22,043). A 2-point increment in MDI score was associated with both all cause mortality (hazard ratio 0.75, 95% CI 0.64-0.87), as well as cause-specific mortality (CHD hazard ratio 0.67, 95% CI 0.47-0.94 and cancer hazard ratio 0.76, 95% CI 0.59-0.98).

As was the case with indexes based on food consumption guidelines, indexes based on established diets are constrained by current knowledge. In other words, they are limited to the diets that researchers have observed and attempted to quantify into a scale or index. They are also limited by our interpretation of what it is within the diet that is worthy of being scored, and our decision regarding where to place cut-off points delineating a healthy level of intake from an unhealthy one.

Dietary intervention studies form another category of diet-based research. These are studies in which participants are randomized to either a specified intervention diet, or to a control diet for a period of time, after which the effects of the dietary intervention are measured. In the Dietary Approaches to Stop Hypertension (DASH) study (Appel et al. 1997), a diet characterized by high fruit, vegetable, and low-fat dairy foods content, together with reduced saturated and total fat content, was compared to both a control diet and an intermediate diet (increased fruits and vegetables, but without the additional attention to fats included in the intervention diet). After an eight week study period, the intervention diet was found to be significantly more successful in lowering both systolic blood pressure (5.5 mm Hg greater reduction) and diastolic blood pressure (3.0 mm Hg greater reduction) than the control diet ($p<0.001$) regardless of baseline hypertension status, although greater reductions in hypertension were observed in study subjects with baseline hypertension. The intermediate diet was also associated with lowered blood pressure levels, however, reductions were not as high as those achieved with the intervention diet.

In the Lyon Heart Study (de Lorgeril et al. 1994), a secondary prevention study for CHD in subjects following myocardial infarction, experimental group subjects were advised to follow a Mediterranean-type diet consisting of increased intake of bread, root and green vegetables, and fish; replacement of beef, lamb, and pork with poultry; daily inclusion of fruit in the diet; and replacement of butter and cream with olive oil and a canola-oil-based margarine supplied by the study. Control subjects were asked to follow a prudent diet. After mean follow-up of 27 months, the probability of survival was

0.82 in the control group and 0.955 in the experimental group (p=0.02), after adjusting for age, gender, smoking, cholesterol, blood pressure, and infarct location; risk ratio 0.30 (p=0.02).

In another CHD secondary prevention study based on the Mediterranean diet and conducted in India, Singh et al. (2002) instructed all study subjects to follow a local diet consistent with National Cholesterol Education Program guidelines, including 30% or less energy from total fat, less than 10% from saturated fat, and less than 300 mg cholesterol per day. In addition, half of the subjects were randomly assigned to the intervention group in which they were additionally advised to consume 400-500 grams of fruits, vegetables, and nuts each day, 400-500 grams of whole grains, legumes, rice, maize, and wheat each day, and 3-4 servings of mustard seed or soy bean oil daily. The intervention diet, considered to be an Indo-Mediterranean diet rich in alpha-linolenic acid, was found to be associated with a significant reduction in sudden cardiac deaths (p=0.015), non-fatal myocardial infarctions (p<0.001), as well as reductions in cardiac risk factors (p<0.001).

Associations with health outcomes

As is evident from the above summary of dietary pattern analyses, many diet/disease relationships have been demonstrated using dietary pattern analysis techniques. Hu (2002) summarized the collective experience with dietary pattern analysis techniques, concluding that they should be used to complement, rather than replace, traditional methods of analysis. "Evidence is enhanced when the results from multiple lines of research (i.e., biomarkers of nutrient intake, nutrients, foods, and dietary patterns) are consistent" (Hu, 2002, p. 8). Hu identifies four circumstances in which dietary pattern analysis may be particularly useful. First, dietary pattern analysis may be helpful in identifying diet/disease relationships when few single dietary associations for disease have been found (as is the case with breast cancer). Second, it may be useful for understanding dietary impact on disease when associations with single aspects of diet have been demonstrated, but overall dietary effects are observed as well (as in CHD). Third, a dietary pattern may be included in multivariate analyses as a covariate along with a specific nutrient, to determine whether the effects of the nutrient alone are independent of the dietary pattern. Fourth, dietary pattern analysis is useful in the evaluation of dietary guidelines.

Rationale for current study

Basiotis et al. (2002) assessed the quality of the American diet by analyzing the NHANES 1999-2000 food intake data against the HEI. In this study, the same dietary intake data (NHANES 1999-2000) are assessed against the MDI. Table 4, Comparison of the Healthy Eating Index (HEI) and the Mediterranean Diet Index (MDI), shows a side-by-side comparison of the two indexes. Both indexes, to a great extent, are based on the intake of foods from broad food group categories, i.e., fruits, vegetables, meats, cereals, etc. However, the HEI also awards 50% of its points on the basis of other criteria, including, percent total fat, percent saturated fat, cholesterol (mg), sodium (mg), and variety. The MDI awards points for two non-food-group components, the ratio of monounsaturated to saturated fat, and moderate intake of alcohol. Furthermore, food group points are scored very differently under the two indexes. The HEI awards points for meeting the minimum serving requirement for each of five food groups on a daily basis, while the Mediterranean Diet Index views consumption of some food groups positively and others negatively, so that greater consumption of some food groups (vegetables, legumes, fruit/nuts, grains and fish) scores points, while limited consumption of other food groups (meat and dairy) scores points. Furthermore, within-food-group differences form another important distinction between the two indexes. For example, under the HEI, potatoes are counted as vegetables. However, in the original application of the index (Trichopoulou et al. 1995) potatoes were counted in the cereals group, while in other applications (Osler & Schroll 1997), they were counted with the vegetables group. In the most recent application of the index (Trichopoulou et al. 2003) potatoes are not incorporated into calculation of the MDI score at all, as they are not generally considered a part of the Mediterranean Diet. The HEI also includes many different protein foods in the meat group (meat, poultry, fish, dry beans, eggs, and nuts), whereas the Mediterranean Diet Index includes only meat and poultry in the meat group, fish and legumes each have their own separate food group categories, nuts are combined with fruits, and eggs are left out altogether (again because eggs are not a regular feature of the Mediterranean Diet) (Trichopoulou et al. 2003). The differentiated manner in which protein-rich foods are scored by the two indexes may impact the ability of each index to measure diet/disease relationships. For example, though fish is an excellent source of protein, it is also a good source of other important nutrients such as omega-3 fatty acids (Hu et al. 2002). The unique nutritional benefits of fish may be lost in the HEI scoring that combines fish consumption with all other protein-rich foods, including meat, which has been shown to be associated with poorer health outcomes (Slattery et al. 1998, Hu et al. 2000). In the Basiotis et al. (2002) assessment of NHANES 1999-2000 against the HEI, it was found that 90% of the US population has a diet that is either poor or is in need of improvement. Analysis of the same data set against the MDI may provide additional insight into the quality and character of the current American diet.

CHAPTER III

METHODS

"The role of statistics is to summarize, to simplify and eventually to explain."

Greenacre, 1984, p. 1

Study population

i. NHANES 1999-2000

The National Health and Nutrition Examination Survey (NHANES) 1999-2000 data set is a complex, multistage, probability sample of the civilian, non-institutionalized population in the United States (NHANES 2003). The data are nationally representative, subject to sampling error. Low-income persons, adolescents 12-19 years old, persons over age 60, African Americans, and Mexican Americans were intentionally over-sampled in the NHANES 1999-2000 survey. Reliable food intake data and demographic data were available for a total of 8,600 respondents from 0 to 85 or more years of age. (Persons 85 or older at the time of the survey interview were recorded as being 85 years old.) Children under 2 years of age (n=530) and pregnant females (n=297) were excluded from the analysis because of their special dietary needs, yielding an analyzable sample of 7,773 individuals.

ii. Subject selection

In this study on dietary quality, all persons age 2 and older were included. While previous investigations using the MDI have been limited to adults, and often the elderly, the analysis of the NHANES 1999-2000 data set by Basiotis et al. (2002) against the HEI included all persons age 2 and above because it is at age 2 that dietary guidelines begin to apply. Furthermore, Basiotis et al. found that young children, ages 2-3, had the highest average HEI score (best quality diet) among all age/gender groups. For these reasons, the same population sample was used in this study; it is possible that this study represents the first application of the MDI score in children.

iii. Collection of food intake data

One of the major objectives of NHANES is "to study the relationship between diet, nutrition, and health" (NHANES 1999-2000, Public Data Release File Documentation). Food intake data were collected by means of 24-hour dietary recalls, recorded using a computer-assisted dietary interview (CADI) system by trained dietary interviewers. The CADI system provided a standardized, scripted interview to facilitate the uniform collection of dietary data from all participants. Proxies (parents or other caregivers) were allowed to provide dietary recall data for participants less than 6 years old, and were allowed to assist in completing the dietary interviews for participants between 6 and 11 years of age. The majority of dietary interviews were conducted in the Mobile Examination Centers (MEC); a sub-sample was conducted by telephone as part of a methodological sub-study evaluating the feasibility of conducting NHANES dietary interviews by telephone. Interviews were conducted in English or Spanish or with the aid of a translator when needed (NHANES 1999-2000, Dietary Interview Component Total Nutrient Intakes File).

vi. Collection of demographic data

Demographic data were collected during an in-home interview using standardized, computer-assisted methods. Responsible adults provided information for participants under 16 years of age, unless there was no such person in the household, in which case persons under 16 were interviewed directly (NHANES 1999-2000, Public Data Release File Documentation).

Statistical analysis

i. Data Management

All analyses were performed with SAS statistical software (SAS Institute, Version 8.0). Graphs were prepared using Excel (Microsoft, 2002). A total of 4,311 unique food items were recorded in the NHANES 1999-2000 survey. For the purposes of this analysis, all food items were categorized into one of seven food groups or were excluded from the food group analysis. Food groups were those defined by the originators of the MDI score and modified to include fish as a separate group in 2003 (Trichopoulou et al. 2003), namely, 1) vegetables (including vegetable juice, but excluding potatoes); 2) legumes (including tofu and soy products); 3) fruit and nuts (including fruit juices); 4) grains (including breads, cereals, crackers, etc., but excluding sweet pastries and desserts); 5) fish (including all seafood); 6) meat (including poultry, but excluding eggs); and 7) dairy (including milk, cream, yogurt and cheese, but excluding butter and ice cream) (Table 4). Food items excluded from the food group analysis were eggs, potatoes, desserts, soft drinks, fruit drinks, non-dairy creamers, condiments and spices, broths, fats and oils, and alcohol. Eggs and potatoes are not major components in the Mediterranean diet, as defined by Trichopoulou et al. (2003), and are therefore not included in MDI scoring. Although it was expected that these two foods contributed a sizeable portion of total energy in the American diet, both foods were excluded in this study in order to consistently apply the MDI score as originally designed, and to facilitate more accurate comparison of American and Greek consumption patterns. Although fats, oils, and alcohol are excluded from the food group analysis, measures of total intake of fat by type (saturated and monounsaturated) and alcohol were collected and included in the MDI score, as described below. Data on polyunsaturated fat intake was analyzed, but was not used in computing the MDI score.

Each of the more than 4,000 food items was reviewed and assigned to one of the seven food groups or the 'excluded' group, and the food intake data for all respondents were categorized accordingly. In addition, approximately 5% of food items were identified as mixtures, in which a significant portion of ingredients came from more than one food group (for example, a hamburger on a bun with lettuce and tomato is a 'mixed' item). Mixtures were divided into their respective contributing food groups according to the weight in grams of the primary ingredients. In the case of the hamburger, for example, the gram weight if the food consumed was divided among the 'meat' group and the 'grains' group in a 50/50 split, (50% meat, 50% grains). All divisions were approximations, determined by assessing the gram weight of the ingredients in the USDA recipes (NTIS, 1998) on which the survey food items were based. Again, in the case of the hamburger, the lettuce and tomato contributed minimally to the weight of the food consumed and consequently were not counted in the mixture. In general, food groups contributing less than 20% of the weight of the item were excluded from the mixture, and the full weight of the item was split among the remaining contributing food groups.

Total intake by food group, measured in grams per day, was calculated for each respondent. Individual food group intake in grams was adjusted to a standardized 2,500 calorie diet for men and boys ages 12 and above, and a 2,000 calorie per day diet for women and girls ages 12 and above. In children 11 years and younger, daily intake was adjusted to a standardized caloric level of 1,300 calories in children ages 2-3, 1,800 calories in children ages 4-6, and 2,000 calories for children ages 7-11. Caloric limits were selected on the basis of previous applications of the MDI in adult populations (Trichopoulou et al. 1995; Osler & Schroll, 1997), and on the standards for children used by Basiotis et al. (2002). In order to calculate the MDI score, median adjusted intake in grams for each food group or nutrient was calculated for each of nine population groups: children 2-3, children 4-6, children 7-11, adolescent females 12-19, adolescent males 12-19, women 20-60, men 20-60, elderly women 61 and older, and elderly men 61 and older. Median intake for each food group was then weighted by multiplying it by the percentage of the population group that consumed items from that food group, to approximate 'usual' intake from each food group.

ii. Determination of Mediterranean Diet Index (MDI) Scores

MDI scores were calculated for each respondent within a population group by comparing adjusted, individual intake to weighted median intake of each food group. If intake was above the median for vegetables, legumes, fruit/nuts, grains, and fish, a score of one point was assessed. If intake was below the median for meat and dairy, a score of one point was assessed. In a similar manner, the median ratio of monounsaturated to saturated fatty acids was calculated for each population group and ratios above the median scored one point. The median ratio of polyunsaturated to saturated fatty acids was calculated for each group as well, but not incorporated into the index score calculations. Alcohol consumption between 10 and 50 grams per day in men (ages 20 and older) and between 5 and 25 grams per day in women (ages 20 and older) also scored one point. MDI scores were summed for each participant; the highest achievable score was 9 for adults and 8 for persons between ages 2 to 19. Following Trichopoulou et al. (2003), scores were grouped into three levels: low, 0-3; moderate, 4-5; and high, 6-8 (age 19 and under) and 6-9 (adults age 20 and older). See Table 4, for a complete listing of MDI score components.

Overall MDI scores for broad demographic groups were compared to HEI analysis results (Basiotis et al. 2002). HEI scores range from 0 to 100. To facilitate comparison of MDI scores with HEI scores, and to adjust for the difference in maximum possible scores for persons age 2-19 and adults age 20 and older, the percent of highest possible MDI score was calculated.

Using the MDI, Trichopoulou et al. (2003) assessed the diets of adults (ages 20-86) in the Greek arm of the EPIC cohort study and reported median intake by food group in this population. Data were collected from 22,043 adults between 1994 and 1999 by means of a food-frequency questionnaire that collected usual intake of 150 food items over a 1 year period. Although the EPIC and NHANES studies differ in design and dietary collection methods, it was thought that a comparison of the dietary profiles resulting from these two studies would nevertheless be instructive. To this end, Greek median food group intake data, as reported by Trichopoulou et al. 2003, were compared to data for American men and women, ages 20 and older.

iii. Measurement of Associations

NHANES 1999-2000 survey data are representative of the United States, non-institutionalized population, and study results may be generalized to that population when factor weights are applied in the analysis. However, application of factor weights was beyond the scope of this study. Therefore, reported results are reflective of the study sample only, and may not be extrapolated to apply to the general United States population.

To simplify the analysis, data for the three separate children's populations were collapsed into a single group, ages 2-11, before testing associations between MDI score and demographic variables. The remaining population groups were not modified, resulting in a total of seven analysis groups. Consistent with most analyses of diet and consistent with Basiotis et al. (2003), the data were analyzed separately by gender in persons age 12 and older. The two age groups that were over-sampled in the survey, 12-19 year olds and adults over age 60, were maintained as intact population groups in this analysis.

Descriptive statistics are provided overall and by population group for MDI score. Weighted median food group intake is provided by population group. In addition, median food group consumption by American adults 20 years and older was compared to that of Greek adults as reported by Trichopoulou et al. (2003). Total energy intake was examined as a possible confounder of MDI score using linear regression analysis. Interaction between birthplace and race/ethnicity was also examined using a general linear model. Associations between MDI score and demographic variables were examined using linear regression analysis by analysis group. Demographic variables included age (relative age within the boundaries of each age-based population group); race/ethnicity (non-Hispanic white, non-Hispanic black, Mexican American, other race, and other Hispanic); birthplace (in or outside of the United States); and education level in adults age 20 and older (less than High School, High School or equivalent, and more than High School).

CHAPTER IV

RESULTS

"The health of the individual and the population in general is the result of interaction between genetics and a number of environmental factors. Nutrition is an environmental factor of major importance."

Simopoulos, 2001, p. 3065S

Mediterranean Diet Index Scores

Demographic data are presented in Table 5, Demographic Characteristics by Analysis Group, and MDI scores, fat ratios, alcohol consumption, and total caloric intake information is presented in Table 6, Mediterranean Diet Index Score, Fat Ratios, Alcohol, and Total Energy by Analysis Group. Data are presented separately and, when applicable, combined for the three groups of children. Data are not presented for education level for those 19 and under, as this information did not apply to the majority of individuals in these analysis groups. An MDI score could not be calculated for one respondent (a male in the 12-19 age group who consumed no food group foods and no fat over the course of the 24-hour recall period), reducing the analyzable sample size of those with MDI scores to 7,772.

Overall, the mean MDI score for the full population was 2.95, with 66.2% of individuals scoring in the 'low' score range (0-3), indicative of poor dietary quality. Only 3.7% had high scores (6 or more), and 30.1% scored in the 'moderate' range (4-5).

MDI scores for the nine population groups are presented in Table 6. Mean MDI scores across all groups were in the upper end of the low score range, extending from a low of 2.76 in the female adolescent group, to a high of 3.13 in elderly men. Adolescents scored the lowest of all population groups, with adolescent males scoring slightly better than adolescent females (Table 6). In addition to having the lowest mean MDI scores, adolescents had the lowest percentages of high and moderate scores (range 27.8-

28.9%). The highest MDI scores were seen at both ends of the age spectrum, with a mean score in elderly men, 61 and older, of 3.13, followed by scores of 3.09 and 3.03 in men 20-60 and children 2-3, respectively. These groups also had the highest percentage of moderate and high scores (range 36.5-39.6%). Only a small minority across all age groups (range 1.4-5.4%) had high scores (6 or higher), while moderate scores (4-5) were received by between a quarter and a third of the population in each of the nine age groups (range 26.4-34.9%).

In children age 2-3 weighted median food group intake (in grams) was highest for the dairy group (299 grams) (Table 7, Median Food Group Intake). The second-highest median intake in this group was from excluded foods (272 grams). In all other population groups the highest gram intake was from excluded foods, ranging from a median 479 grams in children 4-6 years old to a median 1,555 grams in men age 20-60, and comprising from 35 to 65% of daily gram intake. Intake of excluded foods as a proportion of total intake reached its highest point in adult men and women, age 20-60, with 65 and 62% of dietary intake, respectively, and dropped to 54 and 50% of intake in elderly men and women, respectively. Dairy was the second-ranking food group in children ages 4-6, 7-11, and male adolescents age 12-19. Fruit/nuts or grains provided the second-greatest gram intake for female adolescents and all adult groups.

In general, median intake in all food groups increased with age, with the exception of dairy for which the highest median intake was among 4-6 year olds at 342 grams per day, and fruit/nuts, for which the highest median intakes were at either end of the age spectrum in young children and the elderly. Daily gram intake of vegetables increased with age but consistently comprised from 5 to 11% of daily intake, with the highest percentage observed in elderly women. Legumes consistently comprised from 1 to 2% of the daily diet, with highest percentages observed in elderly men. Fish contributed up to 1% of daily intake across population groups with the highest intake observed in elderly men. Meat contributed 6-8% of daily gram intake, with the highest percentages observed in males 12-19. With increased consumption of excluded foods, intake of fruit/nuts, dairy, and to a lesser extent, grains, declined as a percent of total grams consumed per day.

The diet profiles of children closely resembled those of adolescents and adults (Figure 1, Median Intake from the Seven Food Groups). The graphed median values show a fairly consistent curve across age/gender groups, with overall gram intake increasing with age in most groups, but reducing with age in the dairy group. No differences between gender groups were noted, with the exception of generally lower intake, both in grams and in calories, by females compared to males in the same age group. When excluded foods are added to the graph (Figure 2, Median Intake from the Seven Food Groups and Excluded Foods), their relative contribution to total intake is notable. Intake of excluded foods increases with age and soon surpasses intake of food group foods.

The ratio of monounsaturated to saturated fats rose slightly with increasing age from 1.0 in young children, to 1.1 in older children, adolescents, and women 20-60, and then to 1.2 in elderly women and adult men (Table 6). The ratio of polyunsaturated to saturated fats similarly ranged between 0.5 in children and adolescents, to 0.7 in men 20-60 years old.

A total of 82 (6%) women age 20-60, and 36 (5%) women 61 and older received one point toward their total MDI score for moderate alcohol intake. Among men, 223 (18%) age 20-60, and 97 (14%) age 61 and older received one point toward total MDI score as a result of moderate alcohol intake. The majority of adults in all age/gender groups drank no alcohol or too little alcohol to score points in this component of the MDI (Table 6).

A side-by-side comparison of mean HEI scores (Basiotis et al. 2002) and MDI scores for each demographic group is provided in Table 8, Comparison of Mean HEI Scores and Mean MDI Scores in the NHANES 1999-2000 Sample. Overall, there are many similarities between diet assessment results using the two indexes, along with several of notable differences.

Basiotis et al. reported that children age 2-3 had the highest HEI score of any group. In this study as well, this age group had a relatively high mean MDI score of 3.03, however, elderly men age 61 and older and men age 20-60 had higher mean scores. It is likely that the ability for adults to score an extra point as a result of moderate alcohol consumption influenced this outcome, causing the scores of adults and children to be not directly comparable. Calculation of the percent of maximum achievable score allows for direct comparison between age groups, and indicates that the highest score percentages are achieved by children ages 2-11, followed by elderly men, adolescents, adult men, elderly women, and adult women (Table 5). Basiotis et al. saw a decline in dietary quality from early childhood to adolescence, when HEI scores were at their lowest, and then a slow rise back up in older adulthood. MDI scores also declined with increasing age in childhood, with lowest scores in adolescence. When these scores are adjusted by determining the percent maximum possible score, however, it is shown that the lowest score percentage is observed in adult women, followed by an increase in score percentage in the elderly.

Consistent with HEI scores in the Basiotis et al. (2002) study, this study found positive associations between MDI score and both birth outside of the United States and higher education level (Table 8), however, race/ethnicity associations were inconsistent between the two indexes. Basiotis et al. found highest HEI scores in Mexican Americans (64.5), other Hispanics (64.2), and non-Hispanic whites (64.2), while in this study, highest mean MDI scores were observed in the 'other race' group (3.21), the 'other Hispanic' group (3.15), and the 'Mexican American' group (2.98). Non-Hispanic whites had the lowest mean MDI scores (2.89) while non-Hispanic blacks had a slightly higher mean score (2.91). Mean scores for each gender were also a point of difference between the two studies. The HEI index indicated higher dietary quality among females, while the MDI indicated higher dietary quality among males (Table 8).

Comparison of Greek and American food intake (Table 9, Comparative Data for American and Greek Adults) showed that median energy intake was higher in the Greek populations, with Greek women consuming 1,863 calories per day compared to 1,597 calories by American women, and Greek men consuming 2,355 calories per day compared to 2,198 in American men. Total intake (grams) from the seven food groups was higher in Greek adults compared to American adults as well. Greek men consumed 1,440 grams from food groups compared to 949 grams in American men, and Greek women consumed 1,302 grams from food groups compared to 843 grams in American women. Within each country, the profiles were very similar across gender groups, with differences between genders limited to slightly lower gram and caloric intake by women compared to men. However, comparison of median food group intake between Greek adults in this study and American adults in the NHANES 1999-2000 study indicated higher median intake of legumes, grains, and meat by Americans compared to Greeks, lower median intake of fish and dairy, and a considerably lower median intake of vegetables and fruit/nuts by Americans compared to Greeks, generating strikingly different dietary profiles in the two populations (Figure 3, Comparison of Dietary Intake in U.S. and Greek Adults).

In addition to the differences in food group intake between the two populations, a difference in monounsaturated:saturated fat ratio was also apparent (Table 9). The ratios for men and women within each country were identical, but at 1.7 the median ratio for Greek adults was 40% higher than the 1.2 ratio for Americans.

Associations with Demographic Variables

The presence of associations between mean MDI score and demographic variables was assessed using linear regression analysis in the seven population groups. Initial regression analysis indicated that total energy was positively associated with MDI score, therefore, total energy was controlled for in the linear regression model. A significant interaction between birth outside of the United States and race/ethnicity was observed in adult women age 20-60 ($p=0.0012$), but was not observed in any other population group and the interaction term was not incorporated into the regression analysis. Although the population groups are defined by age boundaries, age was retained in the regression equation to test age as a factor associated with MDI score within the population groups. This was particularly important for the broad age range of 20 to 60 in the younger adult group. In addition, the demographic variables race/ethnicity, gender; and birthplace were included in the regression equation for the 2-11 age group; race/ethnicity and birthplace in the equations for each adolescent group (male and female); and race/ethnicity, birthplace, and education level in the regression equations for each of the adult age groups by gender. When race/ethnicity was found to be significantly associated with MDI score in a population group, the analysis was rerun to determine the degree to which each race/ethnic group was associated with the outcome relative to the non-Hispanic white population.

Children 2-11. Age was found to be inversely associated with MDI score in the children's group, age 2-11, with younger children scoring higher than older children ($p=0.0041$) (Table 10, Associations between MDI Score and Demographic Variables). Race/ethnicity, gender, and birthplace were not associated with MDI score in this group.

Females 12-19. In female adolescents, being born outside the United States was positively associated with MDI score ($p=0.0333$), while age and race/ethnicity were not associated with the index score.

Males 12-19. Similarly, in male adolescents birth outside of the United States was associated with a higher MDI score ($p=0.0005$), and in this group race/ethnicity was also significantly associated with score ($p=0.0059$). Sub-analysis indicated that all race/ethnicity groups were significantly associated with higher MDI scores relative to non-Hispanic whites (Table 10).

Women 20-60. In women age 20-60, age ($p<.0001$), birth outside of the United States ($p<0.0001$), and education ($p=0.0032$) were all positively associated with MDI score. Race/ethnicity was not significantly associated with score in this population group, however, a significant interaction between race/ethnicity and birthplace may be masking an association between race/ethnicity and MDI score in this population group.

Men 20-60. Likewise, in men age 20-60, increasing age ($p=0.0004$), birth outside of the United States ($p<0.0001$), and increasing education level ($p<0.0001$) were all positively associated with MDI score, while race/ethnicity was not associated.

Women 61 and older. In women age 61 and older, only race/ethnicity was significantly related to MDI score ($p=0.0098$). Sub-analysis indicated that being non-Hispanic black, Mexican American, or of another race were associated with higher MDI score, while belonging to the 'other Hispanic' group was not associated with higher MDI score (Table 10). The remaining three variables, age, birthplace, and education level were not associated with score in this population group.

Men 61 and older. In men 61 and older, birth outside of the United States and higher education level were significantly associated with MDI score ($p=0.0006$ and $p<0.0001$, respectively). Race/ethnicity was also significantly associated with score (p=0.0151). Sub-analysis indicated that only one race group, 'other race', was significantly associated with higher MDI scores (Table 10). Age was not a significant factor in score in this group.

CHAPTER V

DISCUSSION

"The effect of diet does not occur through the addition of single nutrients; rather, each food combines many nutrients that allow for a synergistic action when present in a certain balance. Moreover, several foods constitute a meal and may reinforce a protective effect or be antagonistic."

Gerber, 2001, p. 3051S

A growing number of nutritional epidemiologists have recognized the importance of studying whole diets and the effects of dietary patterns on health outcomes. The Mediterranean diet has been found to be associated with better health outcomes around the world, both in general populations and in populations with heart disease. In this study of the NHANES 1999-2000 data set, the MDI, an index based on the Mediterranean diet, forms the basis for analysis of dietary quality in the United States. It was hypothesized that the MDI could lend a new perspective on the assessment of dietary quality in an American population.

To a great extent, results in this study are consistent with those reported by Basiotis et al. (2002) who found that 90% of the U.S. population had diets that were either poor or were in need of improvement when measured against the HEI. Similar results were seen in this study, in which the overall, mean MDI score was low at 2.95, indicating an overall low quality diet in the United States; 66.2% of respondents had low MDI scores, 30.1% moderate, and 3.7% had high scores. Analysis by age/gender groups also showed similar results with highest dietary quality observed at either end of the age spectrum, in the very young and the elderly (Table 8). High scores in young children, as Basiotis et al. suggest, may be the result of parents taking extra care to be sure that young children are eating properly. In the elderly, the higher MDI scores may reflect adherence to a more traditional diet, or one that is less based on excluded foods, as it appears that MDI scores rise with a decrease in the proportion of excluded foods in the diet.

When race/ethnicity was a significant factor in MDI score, a significant difference was most often found for 'other race', although significantly higher MDI scores were also found among elderly non-Hispanic black and Mexican American women, and among adolescent males of all race/ethnicity groups other than non-Hispanic whites. The MDI score rewards traditional, plant-based diets. The higher prevalence of such diets may persist in the race/ethnicity groups that are other than non-Hispanic white because these dietary patterns are part of their cultural heritage.

Birth outside of the United States was significantly associated with high MDI score in five of seven population groups. This seems to indicate that many who immigrate to the United States maintain tradition eating patterns after their arrival in this country. Eventually, this high MDI score advantage may disappear if, with successive generations, a more Western diet is adopted (Haenszel & Kurihara, 1968). Immigration to an industrialized country such as the United States is not the only threat to traditional, plant-based, dietary patterns. The same changes may occur within a country in the midst of industrial development. Woo et al. (2001) found higher MDI scores among Chinese living in a rural location in China, than among urban Chinese in Western cities both in and outside of China. He cautioned against China's adoption of negative changes in traditional dietary patterns, such as increased consumption of meat, as the nation develops industrially.

Socio-economic data have not been made available for this data set for reasons of confidentiality; however, education level may be an appropriate proxy for income level. Higher MDI scores were associated with increasing education level in adults. This may result from greater knowledge regarding nutrition, but may also be associated with having greater means to purchase a variety of foods, especially items that contribute positively to MDI score but are more expensive, such as fruit, vegetables, and fish.

The degree to which excluded foods dominate the diet with increasing age, until finally tapering off as a percent of total intake in the elderly, is noteworthy. Excluded foods include both nutritionally poor foods such as desserts, candy and soft drinks, as well as nutritious foods such as eggs and potatoes.

Nevertheless, as consumption of excluded foods increases, relative consumption of more nutritious food group foods declines. Future studies might focus on the analysis of excluded foods, ascertaining which foods these are, and in particular, understanding the proportions of nutrient poor vs. nutritious foods in this category. This analysis was limited to analysis of each food group's contribution (in grams) to total intake. Future studies might alternatively analyze intake of each of the food groups and excluded foods by their contributions of energy (calories) to the diet.

A monounsaturated:saturated fat ratio that is greater than 1.0 and higher than the polyunsaturated:saturated ratio may be viewed as a positive dietary attribute because it means that healthful monounsaturated fats were consumed in greater proportion than both saturated and polyunsaturated fats, although the American ratio of 1.2 falls considerably short of the 1.7 ratio observed in the Greek population. A polyunsaturated:saturated fat ratio less than 1.0 may be viewed as a negative dietary attribute as it indicates greater consumption of saturated than polyunsaturated fat. Though both fat ratios improve with increasing age, reduction in saturated fat consumption would further improve this aspect of the American diet.

Energy intake varied widely within each age/gender group and across study participants, ranging from a low of 6 calories during the 24-hour recall period in one individual, to a high of 9,657 calories in another. However, mean and median total energy levels were closely aligned with the standard caloric intake values used to calculate MDI scores in each of the population groups, with the exception, perhaps of the elderly populations (Table 6). Though standard caloric levels for all adults were adopted on the basis of previous studies (Trichopoulou et al. 1995), in this study, mean energy intake in elderly women was 1,465 calories (median, 1,403 calories) compared to standard intake of 2,000 calories, while in elderly men mean energy intake was 1,913 calories (median, 1,810 calories) compared to standard intake of 2,500 calories (Table 6). The impact of this variance was not analyzed, however, consideration should be given to lowering standard caloric intake levels for the elderly in similar analyses in the future.

Food frequency data from the Greek cohort study and 24-hour recall survey data from the United States national survey both provide summaries of food intake in the adult populations of each country. The patterns of intake by men and women in each country are very similar, while the patterns between countries are noticeably different. Higher energy intake (in calories) by Greek adults may in part explain the greater intake (in grams) from the seven food groups, however, in spite of greater total intake from food groups in the Greek population, Americans had higher median gram intake in three of seven food groups (legumes, grains, and meat), while Greeks had mildly higher intakes of fish and dairy, and considerably higher intakes of fruit/nuts and vegetables.

Median intake values were weighted in this study to approximate usual intake by each populations group. Actual median intake values calculated on the basis of those eating from each food group were multiplied by the percentage of the population group eating from each food group. It is likely that another weighting method might achieve a more accurate measure of usual intake in each population group, and should be considered in future studies of this type. A food-frequency questionnaire pertaining to usual consumption of fish and seafood was newly included in the 1999-2000 version of the NHANES survey, however, these data were not used in this MDI analysis. Incorporation of these data might have resulted in modified median intake values for the fish food group; such an adjustment should also be considered in future studies.

A limitation of this study is the manual categorization of food intake data into food groups, and division of mixtures into their primary food group components. In future studies it would be useful to accomplish this task by means of a computerized process in order to minimize recoding errors. However, in the absence of such tools in this study, the data were carefully coded manually and the USDA recipe database was used to accurately calculate the relative gram weight of ingredients in mixtures. A SAS program was prepared to reliably and uniformly categorize food intake data into one of the seven food groups or the excluded foods group.

Survey sample weights were not applied in this analysis. Without their incorporation, study results may not be generalized to the United States population, however, the current analysis investigates use of the MDI as a tool for assessing dietary quality based on 24-hour recall data from a broad population sample. Study results remain relevant as a description of the study sample. Furthermore, application of sample weights may be carried out at a future time.

Study results are also limited by the reliability of the dietary data. In general, 24-hour recalls collect detailed dietary intake data about a short period of time – one day – while food-frequency questionnaires collect less precise dietary intake data about a longer period of time (Rimm et al. 1992). Because 24-hour recalls collect data on specific foods, they are more precise in measuring nutrient intake and are more appropriate for use in diverse populations, as certain ethnic foods might not be accurately reflected by a general food-frequency questionnaire. Furthermore, 24-hour recalls have been demonstrated to be reliable in measuring the usual intake of populations (Basiotis et al. 2002). For these reasons, the primary instrument used to collect dietary data in national food surveys such as NHANES remains the 24-hour recall (Briefel et al. 1992). Food frequency questionnaires collect usual intake on a limited number of categories of foods (usually about 100-200), and so are useful for assessing dietary patterns and diversity in relation to guidelines (Briefel et al. 1992). Limitations of all dietary data collection instruments, however, are recall bias and underreporting of dietary intake. In this survey, 24-hour recall data are collected by an experienced dietary interviewer using a script that prompts for often-forgotten foods, as a way of addressing the problem of underreporting and creating a uniform data collection environment for all respondents. Furthermore, the period of recall is very brief (24-hours) and in the recent past (the day before the interview), minimizing recall bias.

Overall, the MDI was found to be a useful tool for assessing dietary quality in adults and children. The majority (66%) of the sample studied had low MDI scores, and another 30% had moderate scores. This is consistent with the assessment using the HEI in the same data set that found 16% of the U.S. population to have poor diets, and another 74% to have diets in need of improvement (Basiotis et al. 2002). Other consistencies between the HEI and the MDI assessments were also found, such as higher quality diets among those born outside of the United States, those with higher levels of education, and those in certain subsets of the population, namely the very young and the elderly. The median food group data were useful for graphically characterizing dietary profiles both within the American study sample, and between adults in this study and adults from a separate study conducted in Greece. It was found that the Americans, regardless of age and gender, shared a consistent dietary profile, with increased intake in most food groups increasing with age and associated increases in total energy. It was also shown that, aside from intake of excluded items, the greatest intake in grams per day was from the dairy, fruit/nuts and grains groups. Daily intake of vegetables, legumes, fish and meat remained relatively constant across age/gender groups, while intake of foods excluded from the analysis increased across age/gender groups. The diet profiles of Greeks and Americans differed primarily in the relatively high intake of vegetables and fruit, as well as a 40% higher monounsaturated to saturated fat ratio, in the Greeks compared to the Americans. Future investigations of the MDI in the United States should explore associations between MDI score and long-term health outcomes such as CHD and/or cancer using United States cohort data, the results of which might be compared to similar studies conducted in other parts of the world.

Table 1: *A Posteriori* Analyses of Dietary Patterns

Data	Authors	Year	Method	Association (present/absent/borderline[1])
Ten-State Nutritional Survey, 1968-70, and HANES I, 1971-74	Schwerin et al	1981	PCA	Disease symptoms (present)
Nationwide Food Consumption Survey, 1977-78	Akin et al	1986	Cluster	Sociodemographic factors
Rural, Italian cohorts from Seven Countries Study, 20 year follow-up	Farchi et al	1989	Cluster	Mortality (present)
Bogalusa Heart Study, 1981-83	Nicklas et al	1989	PCA	Cardiovascular risk factors (present)
Western New York Diet Study, 1975-86	Randall et al	1990	PCA	Sociodemographic factors / nutritional adequacy
Elderly Boston-area volunteer survey, 1981-83	Tucker et al	1992	Cluster	Nutritional adequacy
British random sample, 1984-85	Whichelow & Prevost	1996	PCA	Diet, demographic, and lifestyle factors on mortality (present for women)
Framingham Nutrition Studies, 1984-88	Millen et al	1996	Cluster	Nutritional adequacy (Food Guide Pyramid recommendations)
United States urban volunteers survey	Wirfält et al	1997	Cluster	Nutritional adequacy, gender, and weight
Quebec Nutrition Survey, 1990	Beaudry et al	1998	PCA	Nutritional adequacy
Population-based case-control study in 3 states, 1991-94	Slattery et al	1998	PCA	Colon cancer (present)
European Prospective Investigation of Cancer (EPIC) in East Anglia, UK, 1992-94	Fraser et al	2000	Cluster	Demographic factors
Health Professional Follow-up Study, 1986-94	Hu et al	2000	PCA	Coronary heart disease in men (present)
Cross-sectional study in multiethnic women on Hawaii	Maskarinec et al	2000	PCA	Body mass index (present)
National Representative Dietary Survey, 1986-87	Pryer et al	2000	Cluster	Nutritional adequacy and sociodemographic factors
Cross-sectional study in the United Kingdom	Williams et al	2000	PCA	Metabolic syndrome features and lifestyle factors (present)
Health Professional Follow-up Study, 1994	Fung et al	2001	PCA	Biomarkers of obesity and CVD risk (present)
Nurses' Health Study, 1986-94	Fung et al	2001	PCA	Coronary heart disease in women (present)
SU.VI.MAX cohort in France	Guinot et al	2001	MCA	Dietary pattern identification
Framingham Heart (cohort) Study and SENECA (cross-sectional) Study	Haveman-Nies et al	2001	Cluster	Nutritional adequacy and sociodemographic factors
Cohort study in Copenhagen, 1996 follow-up	Osler et al	2001	PCA	Mortality (present)
Swedish mammography screening cohort, 1987-90	Terry et al	2001	PCA	Colorectal cancer (borderline)
Swedish mammography screening cohort, 1987-90	Terry et al	2001	PCA	Breast cancer (absent)

[1]Statistical association between dietary pattern and health outcome either present, absent, or borderline).

Table 1: *A Posteriori* Analyses of Dietary Patterns (continued)

Data	Author	Year	Method	Association (present/absent/borderline[1])
NHANES III, 1988-94	Tseng et al	2001	PCA	Patterns across geographic areas
Malmö Diet and Cancer Cohort, 1991-94	Wirfält et al	2001	Cluster	Metabolic syndrome factors (present)
Framingham Nutrition Studies, 1984-88	Millen et al	2002	Cluster	Carotid atherosclerosis in women (present)
Australian National Nutrition Survey, 1995	Mishra et al	2002	PCA	Socioeconomic status
Cohort study in Copenhagen, 1996 follow-up	Osler et al	2002	PCA	Coronary heart disease (absent after controlling for BMI)
Cross-sectional study in southern France	Siari et al	2002	MCA	Cancer
Framingham Osteoporosis Study, 1948-89	Tucker et al	2002	Cluster	Bone mineral density (present)
Nurses' Health Study, 1986-94	Fung et al	2003	PCA	Colorectal cancer (present)
Whitehall II study of London civil servants, 1991-93	Martikainen et al	2003	Cluster	Socioeconomic factors and gender

[1]Statistical association between dietary pattern and health outcome either present, absent, or borderline).

Table 2: Comparison of Analytical Characteristics of Principal Components Analysis in Three Early Applications

Author	Schwerin et al. 1981	Randall et al. 1990	Slattery et al. 1998
N	32,086	1,475 men 780 women	2,389 men 2,014 women
Data	Cross-sectional	Case-control	Case-control
By gender?	No	Yes	Yes
FA applied to...	15 food groups	110 food items	35 food groups
Eigenvalue	> 1.0	> or = 1.5	> 1.25
Factor Loading	+ or – 0.35	+0.3 and -0.2	+ or – 0.2
Yielded *x* patterns	7	9 men 9 women	6 men 6 women
Variability explained	55.3%	22% men 23.1% women	36.9% men 34.3% women
Patterns	More dairy, soups; less sugary foods, beverages More non-sugary beverages, less dairy More eggs, legumes/nuts, cereal/grains More meats, vegetables/ fruits/juices, desserts More Poultry, less meats More mixed protein, shellfish More fish, fats/oils	**Men** Salad Fruit Staple vegetables Healthful E. European Dessert Fast foods S. European Health foods **Women** Salad S. European/ healthful Fruit Low cost Dessrt Staple vegetables Costly Health foods Non-use	**Men** Western Prudent High fat/sugar dairy Drinker Substituter Fruit juice **Women** Western Prudent High fat/sugar dairy Substituter Coffee and roll Drinker

Table 3: Comparison of Indexes of Dietary Quality

Index Name	Based on	Details
Dietary Diversity Score (DDS)	Five major food groups: dairy, meat, grain, fruit, vegetable	Intake over 24-hr. period. Intake from a food group scores 1 point, for a daily maximum of 5 points. (Kant, *et al.* 1993 and 1995)
Diet Quality Index (DQI)	National Academy of Sciences Publication *Diet and Health* (1989)	Score based on % total fat, % saturated fat, cholesterol (mg), number fruit/vegetable servings, number bread/cereal/legume servings, moderate protein intake, sodium (mg), and adequate calcium (Patterson *et al.* 1994)
Healthy Diet Indicator (HDI)	World Health Organization *Diet, nutrition, and the prevention of chronic diseases* (1990)	Score based on % saturated fat, % polyunsaturated fat, % protein, % complex carbohydrates, dietary fiber (g), fruits/vegetables (g), pulses/nuts/seeds (g), % mono- and disaccharides, and cholesterol (mg) (Huijbregts *et al.* 1997)
Recommended Food Score (RFS)	Combination of dietary guidelines from: National Research Council Committee on Diet and Health (1989), United States Department of Health and Human Services *The Surgeon General's Report on Nutrition and Health* (1988), and USDA *Nutrition and Your Health: Dietary Guidelines for America* (1995)	Measured over a 1 week period. 23 food items from a 62-item food-frequency questionnaire each score 1 point: apples or pears; oranges; cantaloupe; orange or grapefruit juice; grapefruit; other fruit juices; dried beans; tomatoes; broccoli; spinach; mustard, turnip, or collard greens; carrots or mixed vegetables with carrots; green salad; sweet potatoes, yams; other potatoes; baked or stewed chicken or turkey; baked or broiled fish; dark breads; cornbread, tortillas, and grits; high-fiber cereals; cooked cereals; 2% milk and beverages with 2% milk; 1% or skim milk (Kant *et al.* 2000)
Healthy Eating Index (HEI)	USDA *Dietary Guidelines for Americans* (1995)	Score based on servings per day of grains, vegetables, fruit, milk, and meat, as well as on % total fat, % saturated fat, cholesterol (mg), sodium (mg), and variety (McCullough *et al.* 2000a and 2000b)
Mediterranean Dietary Quality Index (DQI)	National Research Council *Diet and Health* (1989) and American Heart Association *Dietary guidelines for healthy American adults* (1996)	Score based on % saturated fat, cholesterol (mg), meat (g), olive oil (mL), fish (g), cereals (g), fruits and vegetables (g), and cigarettes per day (Gerber *et al.* 2000)

Table 4: Comparison of the Healthy Eating Index (HEI) and the Mediterranean Diet Index (MDI)

Healthy Eating Index[1] (daily intake)	Mediterranean Diet Index[2] (daily intake)
Vegetables 3-5 servings	**Vegetables** Intake greater than the median positive (other than potatoes)
	Legumes Intake greater than the median positive
Fruits 2-4 servings	**Fruit and Nuts** Intake greater than the median positive
Grains 6-11 servings (bread, cereal, rice, pasta)	**Grains** Intake greater than the median positive (flour, cereal flakes, starches, pasta, rice, other grains, bread, crisp bread, rusks, breakfast cereals, biscuits, dough, pastry, and other cereal products)
	Fish Intake greater than the median positive
Meat 2-3 servings (meat, poultry, fish, dry beans, eggs, nuts)	**Meat** Intake less than the median positive (meat and poultry, no eggs)
Milk 2-3 servings (milk, yogurt, cheese)	**Dairy** Intake less than the median positive (milk yogurt, cheese)
Total fat intake 30% of energy or less	**Monounsaturated fat:saturated fat** Ratios higher than the median ratio positive
Saturated fat intake Less than 10% of energy	
Cholesterol intake 300 mg or less	**Ethanol** Within prescribed range for gender positive (men: 10-50 g; women: 5-25 g)
Sodium intake 2400 mg or less	
Variety 8 or more different items in a day	
Up to 10 points per category. Maximum score 100. An HEI score over 80 implies a "good" diet, between 51 and 80 a diet that "needs improvement," and 50 or less implies a "poor" diet.	Up to 1 point per category. Maximum score 8 or 9. High score: 6-8 for those 19 or younger and 6-9 for adults due to inclusion of alcohol component in score; moderate score: 4-5; and low score: 0-3.

[1]Basiotis et al. 2002

[2]Trichopoulou et al. 2003

Table 5: Demographic Characteristics by Analysis Group

Age Group	2-3	4-6	7-11	2-11 [1]	Female 12-19	Male 12-19	Female 20-60	Male 20-60	Female 61+	Male 61+
N	370	445	811	1,626	1,055	1,103	1,283	1,268	727	710
Mean Age *(std. dev.)*	2.4 *(0.49)*	5.0 *(0.85)*	9.0 *(1.43)*	6.4 *(2.98*	15.2 *(2.30)*	15.3 *(2.27)*	40.0 *(11.39)*	39.6 *(11.68)*	71.8 *(7.55)*	71.7 *(7.27)*
Race/Ethnicity n *(percent)*										
Non-Hispanic White	106 *(29)*	127 *(21)*	179 *(22)*	412 *(25)*	217 *(21)*	231 *(21)*	505 *(39)*	556 *(44)*	360 *(50)*	369 *(52)*
Non-Hispanic Black	87 *(24)*	108 *(24)*	255 *(31)*	450 *(28)*	298 *(28)*	311 *(28)*	278 *(22)*	255 *(20)*	130 *(18)*	114 *(16)*
Mexican American	134 *(36)*	171 *(38)*	327 *(40)*	632 *(39)*	444 *(42)*	485 *(44)*	370 *(29)*	338 *(27)*	183 *(25)*	180 *(25)*
Other Race (including multiracial)	19 *(5)*	9 *(2)*	12 *(1)*	40 *(2)*	36 *(3)*	25 *(2)*	31 *(2)*	40 *(3)*	12 (2)	14 *(2)*
Other Hispanic	24 *(6)*	30 *(7)*	38 *(5)*	92 *(6)*	60 *(6)*	45 *(4)*	99 *(8)*	82 *(6)*	42 *(6)*	33 *(5)*
Place of Birth n *(percent)*										
United States	355 *(96)*	419 *(94)*	739 *(92)*	1,513 *(94)*	890 *(85)*	883 *(80)*	929 *(73)*	880 *(70)*	542 *(75)*	546 *(77)*
Outside United States	13 *(4)*	25 *(6)*	67 *(8)*	105 *(6)*	158 *(15)*	215 *(20)*	349 *(27)*	375 *(30)*	181 *(25)*	159 *(23)*
Education n *(percent)*										
Less than High School	---	---	---	---	---	---	417 *(33)*	438 *(35)*	355 *(49)*	361 *(51)*
High School (GED)	---	---	---	---	---	---	300 *(23)*	285 *(22)*	172 *(24)*	140 *(20)*
More than High School	---	---	---	---	---	---	563 *(44)*	542 *(43)*	197 *(27)*	205 *(29)*

[1] Group 2-11, used in some analyses, is the combined population from the children's groups 2-3, 4-6, and 7-11.

Table 6: Mediterranean Diet Index Score, Fat Ratios, Alcohol, and Total Energy by Analysis Group

Age Group	2-3	4-6	7-11	2-11[1]	Female 12-19	Male 12-19	Female 20-60	Male 20-60	Female 61+	Male 61+
MDI Score										
Mean *(std.dev.)*	3.03 *(1.39)*	3.0 *(1.36)*	2.91 *(1.33)*	2.96 *(1.36)*	2.76 *(1.30)*	2.77 *(1.33)*	2.79 *(1.35)*	3.09 *(1.48)*	3.0 *(1.37)*	3.13 *(1.43)*
Percent Distribution of MDI Scores										
0-3 (low)	63.5	64.9	67.5	65.9	72.2	71.2	64.5	62.1	66.6	60.4
4-5 (moderate)	31.6	31.5	30.1	30.8	26.4	26.7	30.7	32.5	29.3	34.9
6-8(9) (high)	4.9	3.6	2.5	3.3	1.4	2.2	4.8	5.4	4.1	4.7
Fat Ratios										
Mono:Sat[2]	1.0	1.0	1.1	---	1.1	1.1	1.1	1.2	1.2	1.2
Poly:Sat[3]	0.5	0.5	0.5	---	0.5	0.5	0.6	0.6	0.7	0.6
Alcohol Intake n *(percent)*										
In 'Moderate' Range	---	---	---	---	---	---	82 *(6.4)*	223 *(17.6)*	36 *(5.0)*	97 *(13.7)*
Above Range	---	---	---	---	---	---	100 *(7.8)*	141 *(11.1)*	28 *(3.9)*	35 *(4.9)*
Below Range	---	---	---	---	---	---	1,101 *(85.8)*	904 *(71.3)*	663 *(91.2)*	578 *(81.4)*
Total Energy in Calories										
Mean *(std. dev.)*	1,543 *(631.9)*	1,733 *(614.2)*	2,013 *(812.0)*	1,829 *(748.7)*	1,907 *(888.5)*	2,486 *(1163.1)*	1,846 *(804.0)*	2,618 *(1167.2)*	1,465 *(597.2)*	1,913 *(773.2)*
Median	1,441	1,630	1,888	1,698	1,801	2,279	1,708	2,420	1,403	1,819
Minimum	170	537	367	170	119	20	99	201	12	6
Maximum	5,062	4,367	5,900	5,900	9,606	9,657	8,361	8,935	4,334	5,333
Standard Caloric Intake for Age/Gender Group	1,300	1,800	2,000	---	2,000	2,500	2,000	2,500	2,000	2,500

[1]Group 2-11, used in some analyses, is the combined population from the children's groups 2-3, 4-6, and 7-11. [2]Ratio of monounsaturated:saturated fat. [3]Ratio of polyunsaturated:saturated fat.

Table 7: Median Food Group Intake

Age Group	2-3	4-6	7-11	Female 12-19	Male 12-19	Female 20-60	Male 20-60	Female 61+	Male 61+
Median Intake in Grams									
Excluded	271.5 *(26)*	479.1 *(35)*	611.8 *(41)*	885.2 *(55)*	1,098.2 *(55)*	1,215.0 *(62)*	1,554.8 *(65)*	992.6 *(50)*	1,350.4 *(54)*
Vegetables	54.1 *(5)*	77.8 *(6)*	81.1 *(5)*	92.4 *(6)*	94.1 *(5)*	151.1 *(8)*	152.1 *(6)*	211.3 *(11)*	213.3 *(9)*
Legumes	14.8 *(1)*	13.1 *(1)*	15.3 *(1)*	16.0 *(1)*	20.1 *(1)*	24.4 *(1)*	35.2 *(1)*	26.3 *(1)*	48.3 *(2)*
Fruit/Nuts	204.5 *(20)*	196.3 *(14)*	151.6 *(10)*	133.7 *(8)*	119.1 *(6)*	127.7 *(6)*	120.4 *(5)*	236.2 *(12)*	236.0 *(10)*
Grains	123.1 *(12)*	189.8 *(14)*	200.7 *(14)*	188.2 *(12)*	230.8 *(12)*	178.7 *(9)*	221.8 *(9)*	200.3 *(10)*	225.6 *(9)*
Fish	3.2 *(-)*	2.9 *(-)*	5.4 *(-)*	7.6 *(-)*	7.2 *(-)*	16.7 *(1)*	18.3 *(1)*	16.2 *(1)*	22.8 *(1)*
Meat	58.1 *(6)*	85.4 *(6)*	99.8 *(7)*	101.8 *(6)*	149.5 *(8)*	117.9 *(6)*	162.6 *(7)*	115.3 *(6)*	171.8 *(7)*
Dairy	299.1 *(29)*	341.9 *(25)*	317.0 *(21)*	182.4 *(11)*	263.3 *(13)*	138.3 *(7)*	141.3 *(6)*	200.1 *(10)*	216.4 *(9)*

Table 8: Comparison of Mean HEI and Mean MDI Scores in the NHANES 1999-2000 Sample

	Mean HEI Score[1]	Mean MDI Score
Gender		
Male	63.2	2.99 *(35.1)*[2]
Female	64.5	2.91 *(34.1)*
Age/Gender		
Children 2-3	75.7	3.03 *(37.9)*
Children 4-6	66.9	3.0 *(37.5)*
Children 7-10	66.0	---
Children 7-11	---	2.91 *(36.4)*
Females 11-14	61.4	---
Females 15-18	61.7	---
Females 12-19	---	2.76 *(34.5)*
Males 11-14	60.8	---
Males 15-18	59.9	---
Males 12-19	---	2.77 *(34.6)*
Females 19-50	63.2	---
Females 20-60	---	2.79 *(31.0)*
Males 19-50	61.3	---
Males 20-60	---	3.09 *(34.3)*
Females 51+	66.6	---
Females 61+	---	3.0 *(33.3)*
Males 51+	65.2	---
Males 61+	---	3.13 *(34.8)*
Race/Ethnicity		
Non-Hispanic White	64.2	2.89 *(33.9)*
Non-Hispanic Black	61.1	2.91 *(34.1)*
Mexican American	64.5	2.98 *(35.0)*
Other Race	63.4	3.21 *(37.6)*
Other Hispanic	64.2	3.15 *(36.9)*
Birthplace		
United States	63.5	2.88 *(33.8)*
Mexico	66.0	3.17 *(37.2)*
Other	65.7	3.4 *(39.9)*
Education[3]		
< High School	61.1	2.98 *(35.0)*
High School	63.0	2.86 *(33.5)*
> High School	65.3	3.23 *(37.9)*

[1]Basiotis et al. 2002
[2]Percent of total maximum score. Maximum MDI score was '8' in populations age 2-19, '9' in populations age 20 and older, and '8.526' in the overall population of 3,783 age 2-19 and 3,988 age 20 and older.
[3]HEI scores measured in persons age 25 and older; MDI scores measured in persons age 20 and older.

Table 9: Comparative Data for American and Greek Adults[1]

Group	U.S. Women[2]	U.S. Men[2]	Greek Women	Greek Men
N	2,010	1,978	13,148	8,895
Median Energy Intake in Calories	1,597	2,198	1,863	2,355
Mono:Sat Fat Ratio[3]	1.2	1.2	1.7	1.7
Median Intake in Grams *(percent)*				
Vegetables	169.2 *(20)*	170.2 *(18)*	499.6 *(38)*	549.9 *(38)*
Legumes	23.9 *(3)*	40.7 *(4)*	6.7 *(1)*	9.1 *(1)*
Fruit/Nuts	167.9 *(20)*	157.8 *(17)*	356.3 *(27)*	362.5 *(25)*
Grains	185.3 *(22)*	223.0 *(24)*	139.7 *(11)*	177.7 *(12)*
Fish	16.2 *(2)*	19.4 *(2)*	18.8 *(1)*	23.7 *(2)*
Meat	117.6 *(14)*	166.0 *(18)*	89.8 *(7)*	120.8 *(8)*
Dairy	162.9 *(19)*	171.6 *(18)*	191.1 *(15)*	196.7 *(14)*
Total Median Intake in Grams from Food Groups	843.0	948.6	1,302.0	1,440.4

[1]Greek data from Trichopoulou et al. 2003.

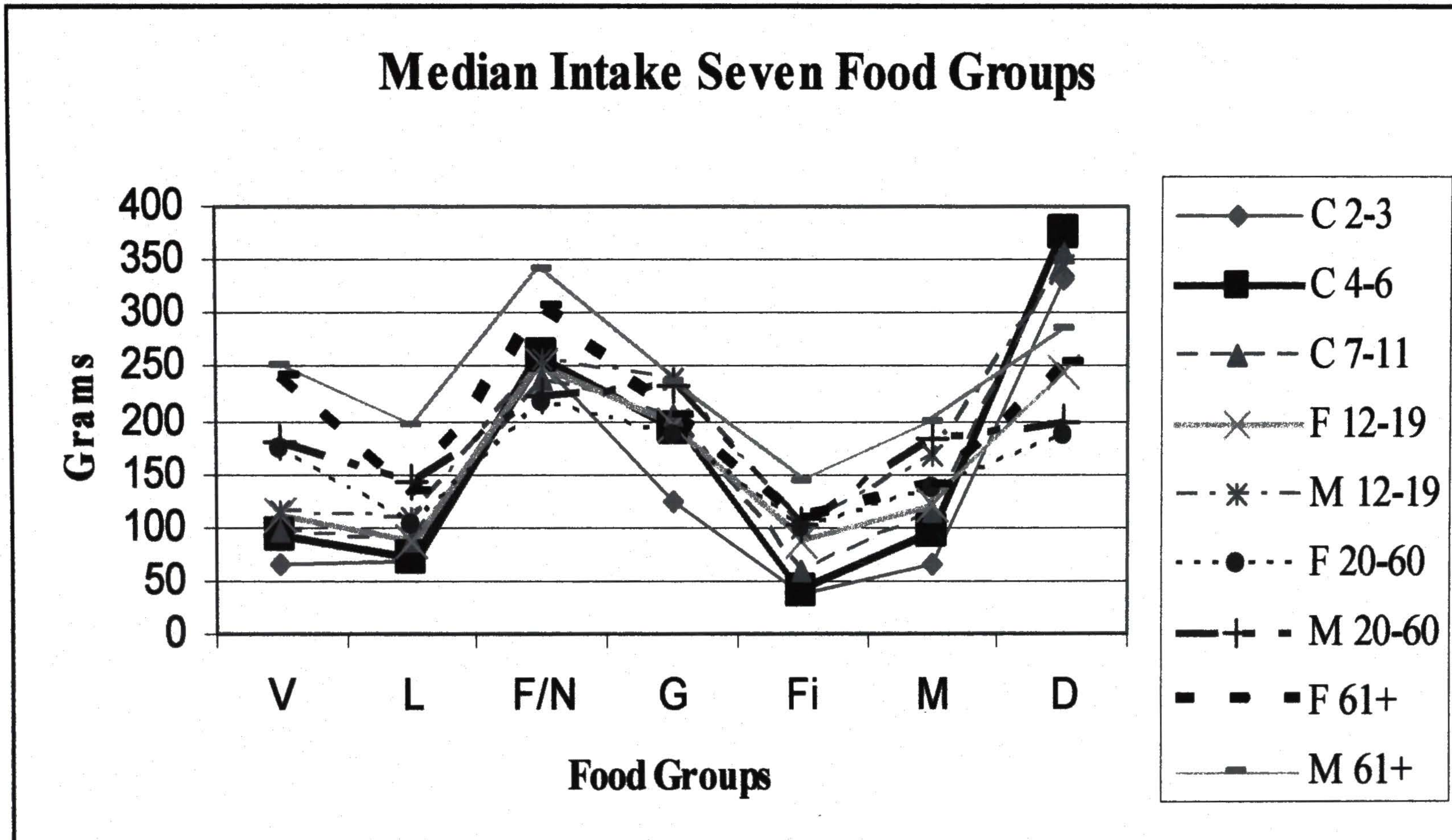
[2]Data for American men and women are the combined data from the two analysis groups for each gender, ages 20-60, and ages 61 and older.

[3]Ratio of monounsaturated:saturated fat

Table 10: Associations between MDI Score and Demographic Variables

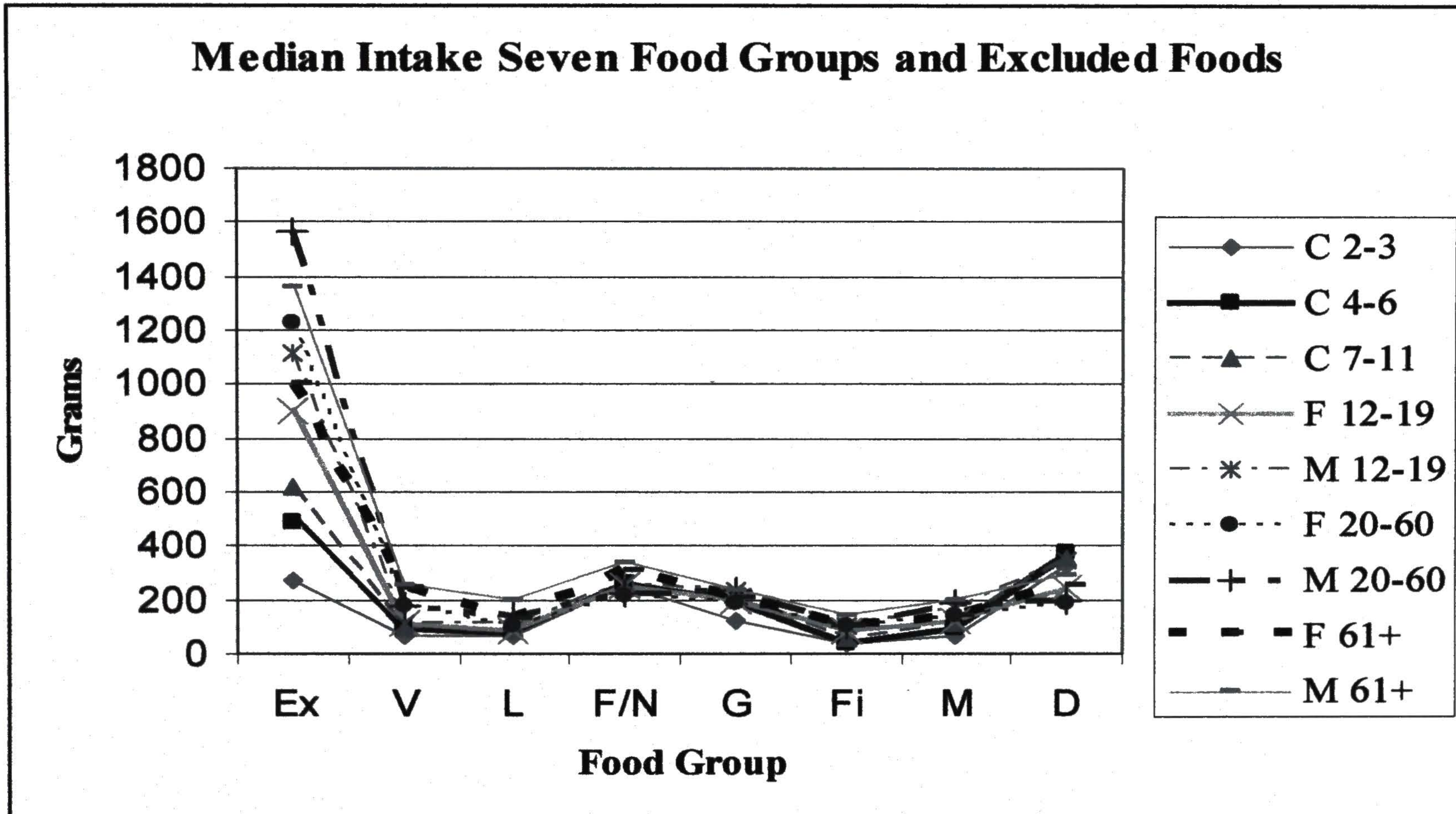
Age Group	2-11	Female 12-19	Male 12-19	Female 20-60	Male 20-60	Female 61+	Male 61+
Demographic Variables p *(direction of significant association)*							
Age	0.0041 (-)	0.2405	0.6079	<0.0001 (+)	0.0004 (+)	0.7000	0.4691
Gender	0.9662	---	---	---	---	---	---
Race/Ethnicity	0.8347	0.3656	0.0059	0.3463	0.5596	0.0098	0.0151
Non-Hispanic Black			(+)			(+)	
Mexican American			(+)			(+)	
Other Race			(+)			(+)	(+)
Other Hispanic			(+)				
Birth Outside U.S.	0.9022	0.0333 (+)	0.0085 (+)	<0.0001 (+)	<0.0001 (+)	0.0940	0.0006 (+)
Education Level	---	---	---	0.0032 (+)	<0.0001 (+)	0.5148	<0.0001 (+)

Figure 1: Median Intake from the Seven Food Groups



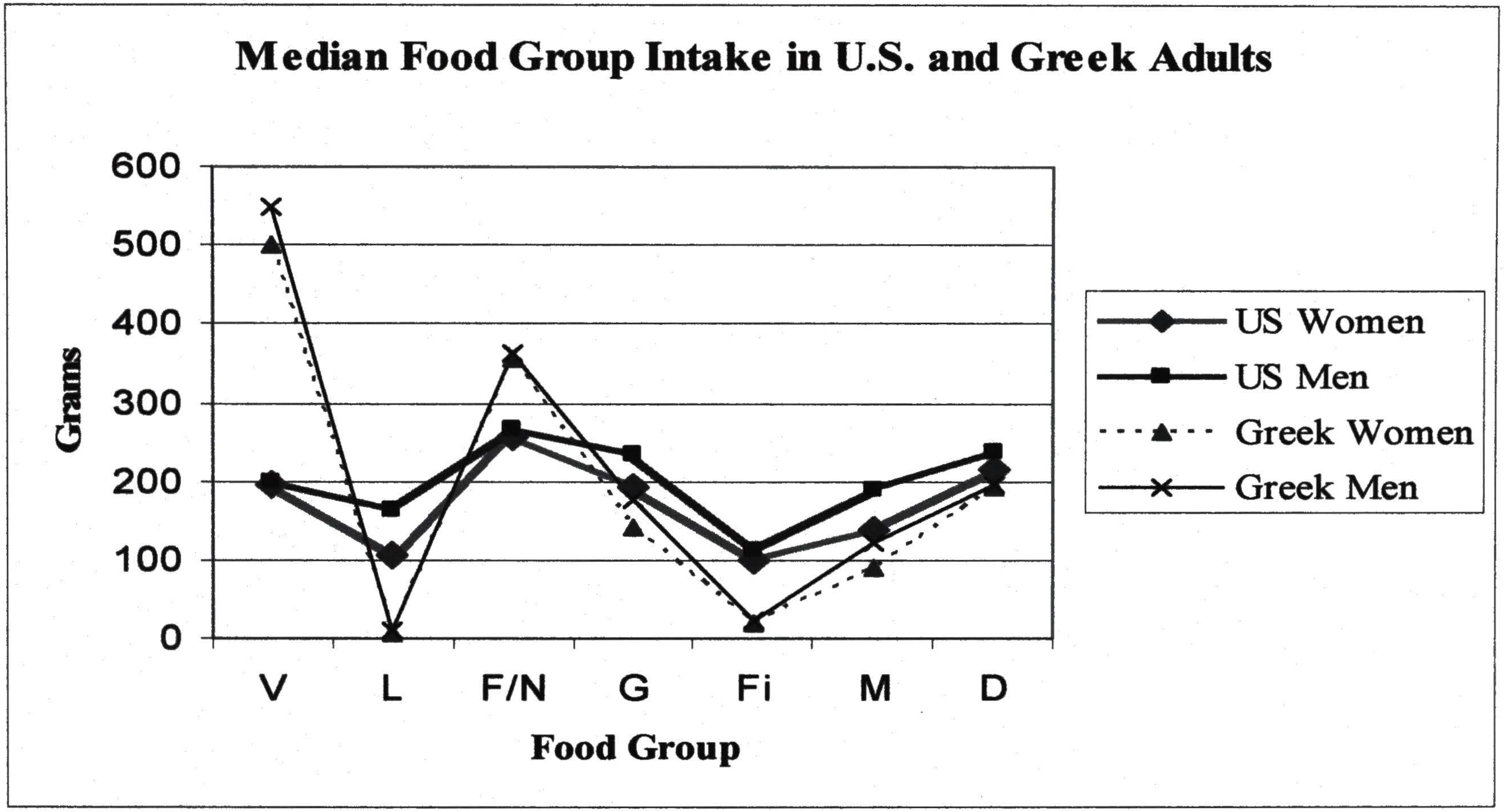
Explanation of Abbreviations: V – Vegetables; L – Legumes; F/N – Fruit/Nuts; G – Grains; Fi – Fish; M – Meat; D – Dairy; C2-3 – Children Age 2-3; C4-6 – Children Age 4-6; C 7-11 – Children Age 7-11; F 12-19 – Females Age 12-19; M 12-19 – Males 12-19; F 20-60 – Females 20-60; M 20-60 – Males 20-60; F 61+ - Females 61 and Older; M 61+ - Males 61 and Older

Figure 2: Median Intake from the Seven Food Groups and Excluded Foods



Explanation of Abbreviations: V – Vegetables; L – Legumes; F/N – Fruit/Nuts; G – Grains; Fi – Fish; M – Meat; D – Dairy; C2-3 – Children Age 2-3; C4-6 – Children Age 4-6; C 7-11 – Children Age 7-11; F 12-19 – Females Age 12-19; M 12-19 – Males 12-19; F 20-60 – Females 20-60; M 20-60 – Males 20-60; F 61+ - Females 61 and Older; M 61+ - Males 61 and Older

Figure 3: Comparison of Dietary Intake in U.S. and Greek Adults



Explanation of Abbreviations: V – Vegetables; L – Legumes; F/N – Fruit/Nuts; G – Grains; Fi – Fish; M – Meat; D – Dairy; C2-3 – Children Age 2-3; C4-6 – Children Age 4-6; C 7-11 – Children Age 7-11; F 12-19 – Females Age 12-19; M 12-19 – Males 12-19; F 20-60 – Females 20-60; M 20-60 – Males 20-60; F 61+ - Females 61 and Older; M 61+ - Males 61 and Older

REFERENCE LIST

Akin, J. S., Guilkey, D. K., Popkin, B. M., & Fanelli, M. T. (1986) Cluster analysis of food consumption patterns of older Americans. *Journal of the American Dietetic Association, 86,* 616-624.

Appel, L. J., Moore, T. J., Obarzanek, E., Vollmer, W. M., Svetkey, L. P., Sacks, F. M., et al. (for the DASH Collaborative Research Group). (1997) A clinical trial of the effects of dietary patterns on blood pressure. *New England Journal of Medicine, 336,* 1117-1124.

Armitage, P., Berry, G., & Matthews, J.N.S. (2002) *Statistical Methods in Medical Research* (4th ed.). London: Blackwell Science.

Basiotis, P. P., Carlson, A., Gerrior, S. A., Juan, W. Y., & Lino, M. (Authors in alphabetical order.) (2002) *The Healthy Eating Index: 1999-2000.* Washington, DC: United States Department of Agriculture, Center for Nutrition Policy and Promotion. (CNPP-12.)

Beaudry, M., Galibois, I., & Chaumette, P. (1998) Dietary patterns of adults in Québec and their nutritional adequacy. *Canadian Journal of Public Health; 89,* 347-351.

Carroll, K. K. (1975) Experimental evidence of dietary factors and hormone-dependent cancers. *Cancer Research; 35,* 3374-3383.

Cumming, R. G., Cummings, S. R., Nevitt, M. C., Scott, J., Ensrud, K. E., Vogt, T. M., et al. (1997) Calcium intake and fracture risk: Results from the study of osteoporotic fractures. *American Journal of Epidemiology, 145,* 926-934.

Cumming, R. G., & Nevitt, M. C. (1997) Calcium for prevention of osteoporotic fractures in postmenopausal women. *Journal of Bone and Mineral Research, 12,* 1321-1329.

Farchi, G., Mariotti, S., Menotti, A., Seccareccia, F., Torsello, S., & Fidanza, F. (1989) Diet and 20-y mortality in two rural population groups of middle-aged men in Italy. *American Journal of Clinical Nutrition, 50,* 1095-1103.

Feskanich, D., Willett, W. C., Stampfer, M. J., & Colditz, G. A. (1997) Milk, dietary calcium, and bone fractures in women: A 12-year prospective study. *American Journal of Public Health, 87,* 992-997.

Fraser, G. E., Welch, A., Luben, R., Bingham, S. A., & Day, N. E. (2000) The effect of age, sex, and education on food consumption of a middle-aged English cohort – EPIC in East Anglia. *Preventive Medicine, 30,* 26-34.

Fung, T., Hu, F. B., Fuchs, C., Giovannucci, E., Hunter, D. J., Stampfer, M. J., et al. (2003) Major dietary patterns and the risk of colorectal cancer in women. *Archives of Internal Medicine, 163,* 309-314.

Fung, T. T., Rimm, E. B., Spiegelman, D., Rifai, N., Tofler, G. H., Willett, W. C., et al. (2001) Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *American Journal of Clinical Nutrition, 73,* 61-67.

Fung, T. T., Willett, W. C., Stampfer, M. J., Manson, J. E., & Hu, F. B. (2001) Dietary patterns and the risk of coronary heart disease in women. *Archives of Internal Medicine, 161,* 1857-1862.

Gerber, M. (2001) The comprehensive approach to diet: A critical review. *Journal of Nutrition, 131(Suppl.),* 3051S-3055S.

Gerber, M. J., Scali, J. D., Michaud, A., Durand, M. D., Astre, C. M., Dallongeville, J., et al. (2000) Profiles of a healthful diet and its relationship to biomarkers in a population sample from Mediterranean southern France. *Journal of the American Dietetic Association, 100,* 1164-1171.

Goodwyn, P. J., & Boyd, N. F. (1987) Critical appraisal of the evidence that dietary fat intake is related to breast cancer risk in humans. *Journal of the National Cancer Institute; 79,* 473-485.

Greenacre, M.J. (1984) *Theory and Applications of Correspondence Analysis.* London: Academic Press.

Greenwald, P. (1999) Role of dietary fat in the causation of breast cancer: Point. *Cancer Epidemiology, Biomarkers & Prevention, 8,* 3-7.

Guinot, C., Latreille, J., Malvy, D., Preziosi, P., Galan, P., Hercberg, S., et al. (2001) Use of multiple correspondence analysis and cluster analysis to study dietary behaviour: Food consumption questionnaire in the SU.VI.MAX. cohort. *European Journal of Epidemiology, 17,* 505-516.

Haenszel, W. & Kurihara, M. (1968) Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. *Journal of the National Cancer Institute, 40,* 43-68.

Haveman-Nies, A., Tucker, K. L., de Groot, L. C. P. G. M., Wilson, P. W. F., & van Staveren, W. A. (2001) Evaluation of dietary quality in relationship to nutritional and lifestyle factors in elderly people of the US Framingham Heart Study and the European SENECA study. *European Journal of Clinical Nutrition, 55,* 870-880.

Hu, F. B. (2002) Dietary pattern analysis: A new direction in nutritional epidemiology. *Current Opinions in Lipidology, 13,* 3-9.

Hu, F. B., Bronner, L., Willett, W. C., Stampfer, M. J., Rexrode, K. M., Albert, C. M., et al. (2002) Fish and Omega-3 fatty acid intake and risk of coronary heart disease in women. *Journal of the American Medical Association, 287,* 1815-1821.

Hu, F. B., Rimm, E., Smith-Warner, S. A., Feskanich, D., Stampfer, M. J., Ascherio, A., et al. (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *American Journal of Clinical Nutrition, 69,* 243-249.

Hu, F. B., Rimm, E. B., Stampfer, M. J., Ascherio, A., Spiegelman, D., & Willett, W. C. (2000) Prospective study of major dietary patterns and risk of coronary heart disease in men. *American Journal of Clinical Nutrition, 72,* 912-921.

Hu, F. B., Stampfer, M. J., Manson, J. E., Rimm, E., Colditz, G. A., Rosner, B. A., et al. (1997) Dietary fat intake and the risk of coronary disease in women. *New England Journal of Medicine, 337,* 1491-1499.

Huijbregts, P., Feskens, E., Räsänen, L., Fidanza, F., Nissinen, A., Menotti, A., et al. (1997) Dietary pattern and 20 year mortality in elderly men in Finland, Italy, and the Netherlands: Longitudinal cohort study. *BMJ, 315,* 13-17.

Hunter, D. J., Spiegelman, D., Adami, H.-O., Beeson, L., van den Brandt, P. A., Folsom, A. R., et al. (1996) Cohort studies of fat intake and the risk of breast cancer - a pooled analysis. *New England Journal of Medicine, 334,* 356-361.

Hunter, D. J. (1999) Role of dietary fat in the causation of breast cancer: Counterpoint. *Cancer Epidemiology, Biomarkers & Prevention, 8,* 9-13.

Jacques, P. F., & Tucker, K. L. (2001) Are dietary patterns useful for understanding the role of diet in chronic disease? *American Journal of Clinical Nutrition, 73,* 1-2.

Kant, A. K. (1996) Indexes of overall diet quality: A review. *Journal of the American Dietetic Association, 96,* 785-791.

Kant, A. K., Schatzkin, A., Graubard, B. I., & Schairer, C. (2000) A prospective study of diet quality and mortality in women. *Journal of the American Medical Association, 283,* 2109-2115.

Kant, A. K., Schatzkin, A., Harris, T. B., Ziegler, R. G., & Block, G. (1993) Dietary diversity and subsequent mortality in the First National Health and Nutrition Examination Survey Epidemiologic Follow-up Study. *American Journal of Clinical Nutrition, 57,* 434-440.

Kant, A. K., Schatzkin, A., & Ziegler, R. G. (1995) Dietary diversity and subsequent cause-specific mortality in the NHANES I Epidemiologic Follow-up Study. *Journal of the American College of Nutrition, 14,* 233-238.

Keys, A. (1980) *Seven countries: A multivariate analysis of death and coronary heart disease.* Cambridge, MA: Harvard University Press.

Keys, A., Aravanis, C., van Buchem, F. S. P., Blackburn, H., Buzima, R., Djordjević, B. S., et al. (1981) The diet and all-causes death rate in the seven countries study. *Lancet, 2,* 58-61.

Kouris-Blazos, A., Gnardellis, C., Wahlqvist, M. L., Trichopoulos, D., Lukito, W., & Trichopoulou, A. (1999) Are the advantages of the Mediterranean diet transferable to other populations? A cohort study in Melbourne, Australia. *British Journal of Nutrition, 82,* 57-61.

Kuh, D., & Ben-Shlomo, Y. (1997) Conclusions. In D. Kuh D & Y. Ben-Shlomo (Eds.), *A life course approach to chronic disease epidemiology* (pp. 296-307). New York: Oxford University Press.

Kuh, D., Power, C., Blane, D., & Bartley, M. (1997) Social pathways between childhood and adult health. In D. Kuh D & Y. Ben-Shlomo (Eds.), *A life course approach to chronic disease epidemiology* (pp. 169-198). New York: Oxford University Press.

Kushi, L. H., & Foerster, S. B. (1998) Diet and nutrition. In R. C. Brownson, P. L. Remington, & J. R. Davis (Eds.), *Chronic disease epidemiology and control* (2nd ed., pp. 215-259). Washington, DC: American Public Health Association.

Kushi, L. H., & Kottke, T. E. (1990) Dietary fat and coronary heart disease: Evidence for a causal relation. In R. B. Goldbloom & R. S. Lawrence (Eds.), *Preventing disease: Beyond the rhetoric* (pp. 385-400). New York: Springer.

Lasheras, C., Fernandez, S., & Patterson, A. M. (2000) Mediterranean diet and age with respect to overall survival in institutionalized, nonsmoking elderly people. *American Journal of Clinical Nutrition, 71,* 987-992.

de Lorgeril, M., Renaud, S., Mamelle, N., Salen, P., Martin, J.-L., Monjaud, I., et al. (1994) Mediterranean alpha-linolenic acid-rich diet in secondary prevention of coronary heart disease. *Lancet, 343,* 1454-1459.

de Lorgeril, M., Salen, P., Martin, J.-L., Monjaud, I., Delaye, J., Mamelle, N. (1999) Mediterranean diet, traditional risk factors, and the rate of cardiovascular complications after myocardial infarction. Final report of the Lyon Diet Heart Study. *Circulation, 99,* 779-785.

Martikainen, P., Brunner, E., & Marmot, M. (2003) Socioeconomic differences in dietary patterns among middle-aged men and women. *Social Science & Medicine, 56,* 1397-1410.

Martínez, M. E., Marshall, J. R., & Sechrest, L. (1998) Factor analysis and the search for objectivity [Invited commentary]. *American Journal of Epidemiology, 148,* 17-19.

Maskarinec, G., Novotny, R., & Tasaki, K. (2000) Dietary patterns are associated with body mass index in multiethnic women. *The Journal of Nutrition, 130,* 3068-3072.

McCullough, M. L., Feskanich, D., Rimm, E. B., Giovannucci, E. L., Ascherio, A., Variyam, J. N., et al. (2000) Adherence to the *Dietary Guidelines for Americans* and risk of major chronic disease in men. *American Journal of Clinical Nutrition, 72,* 1223-1231.

McCullough, M. L., Feskanich, D., Stampfer, M. J., Rosner, B. A., Hu, F. B., Hunter, D. J., et al. (2000) Adherence to the *Dietary Guidelines for Americans* and risk of chronic disease in women. *American Journal of Clinical Nutrition, 72,* 1214-1222.

McKenna, M. T., Taylor, W. R., Marks, J. S., & Koplan, J. P. (1998) Current issues and challenges in chronic disease control. In R. C. Brownson, P. L. Remington, & J. R. Davis (Eds.), *Chronic disease epidemiology and control* (2nd ed., pp. 1 - 26). Washington, DC: American Public Health Association.

Millen, B. E., Quatromoni, P. A., Copenhafer, D. L., Demissie, S., O'Horo, C. E., & D'Agostino, R. B. (2001) Validation of a dietary pattern approach for evaluating nutritional risk: The Framingham Nutrition Studies. *Journal of the American Dietetic Association, 101,* 187-194.

Millen, B. E., Quatromoni, P. A., Gagnon, D. R., Cupples, L. A., Franz, M. M., & D'Agostino, R. B. (1996) Dietary patterns of men and women suggest targets for health promotion: The Framingham Nutrition Studies. *American Journal of Health Promotion, 11,* 42-53.

Milton, R. C. (1998) Quantiles. In P. Armitage & T. Colton (Eds.). *Encyclopaedia of biostatistics* (pp. 3328-3329). Chichester, UK: Wiley.

Mishra, G., Ball, K., Arbuckle, J., & Crawford, D. (2002) Dietary patterns of Australian adults and their association with socioeconomic status: Results from the 1995 National Nutrition Survey. *European Journal of Clinical Nutrition, 56,* 687-693.

NHANES National Health and Examination Survey – NHANES 1999-2000 Data Files. Study documentation including 'Public Data Release File Documentation' and 'Addendum to the NHANES III Analytic Guidelines' and 'Dietary Interview Component Total Nutrient Intakes File' retrieved June 27, 2003, from CDC National Center for Health Statistics Web site: http://www.cdc.gov/nchs/about/major/nhanes/NHANES99_00.htm

NTIS National Technical Information Service. (1998). Continuing Survey of Food Intakes by Individuals (CSFII 1994-96 CD-ROM with Search and Retrieval Software). Data file. (NTIS Order Number PB2000-500027) Beltsville, MD: Agricultural Research Service, U.S. Department of Agriculture.

Nicklas, T. A., Webber, L. S., Thompson, B., & Berenson, G. S. (1989) A multivariate model for assessing eating patterns and their relationship to cardiovascular risk factors: The Bogalusa Heart Study. *American Journal of Clinical Nutrition, 49,* 1320-1327.

Nube, M., Kok, F. J., Vandenbrouke, J. P., van der Heide-Wessel, C., & van der Heide, R. M. (1987) Scoring of prudent dietary habits and its relation to 25-year survival. *Journal of the American Dietetic Association, 87,* 171-175.

Osler, M., Andreasen, A. H., Heitmann, B., Høidrup, S., Gerdes, U., Jørgensen, L. M., et al. (2002) Food intake patterns and risk of coronary heart disease: A prospective cohort study examining the use of traditional scoring techniques. *European Journal of Clinical Nutrition, 56,* 568-574.

Osler, M., Heitmann, B. L., Gerdes, L. U., Jørgensen, L. M., & Schroll, M. (2001) Dietary patterns and mortality in Danish men and women: A prospective observational study. *British Journal of Nutrition, 85,* 219-225.

Osler, M., & Schroll, M. (1997) Diet and mortality in a cohort of elderly people in a North European community. *International Journal of Epidemiology, 26,* 155-159.

Patterson, R. E., Haines, P. S., & Popkin, B. M. (1994) Diet quality index: Capturing a multidimensional behavior. *Journal of the American Dietetic Association, 94,* 57-64.

Perry, I. J. (1997) Fetal growth and development: The role of nutrition and other factors. In D. Kuh D & Y. Ben-Shlomo (Eds.), *A life course approach to chronic disease epidemiology* (pp. 145-168). New York: Oxford University Press.

Pryer, J. A., Nichols, R., Elliott, P., Thakrar, B., Brunner, E., & Marmot, M. (2001) Dietary patterns among a national random sample of British adults. *Journal of Epidemiology & Community Health, 55,* 29-37.

Quatromoni, P. A., Copenhafer, D. L., Demissie, S., D'Agostino, R. B., O'Horo, C. E., Nam, B.-H., et al. (2002) The internal validity of a dietary pattern analysis. The Framingham Nutrition Studies. *Journal of Epidemiology & Community Health, 56,* 381-388.

Randall, E., Marshall, J. R., Graham, S., & Brasure, J. (1990) Patterns in food use and their associations with nutrient intakes. *American Journal of Clinical Nutrition, 52,* 739-745.

Reid, I. R., Ames, R. W., Evans, M. C., Gamble, G. D., & Sharpe, S. J. (1993) Effect of calcium supplementation on bone loss in postmenopausal women. *New England Journal of Medicine, 328,* 460-464.

Schwerin, H. S., Stanton, J. L., Riley, A. M., Jr., Schaefer, A. E., Leveille, G. A., Elliott, J. G., et al. (1981) Food eating patterns and health: A reexamination of the Ten-State and HANES I surveys. *American Journal of Clinical Nutrition, 34,* 568-580.

Shekelle, R. B., Shyrock, A. M., Paul, O., Lepper, M., Stampfer, J., Liu, S., et al. (1981) Diet, serum cholesterol, and death from coronary heart disease: The Western Electric Study. *New England Journal of Medicine, 304,* 65-70.

Siari, S., Scali, J., Richard, A., Tretarre, B., Daurès, J. P., Padilla, M., et al. (2002) Subregional variations of dietary consumption and incidences of cancers in southern France. *IARC Scientific Publications, 156,* 127-129.

Singh, R. B., Dubnov, G., Niaz, M. A., Ghosh, S., Singh, R., Rastogi, S. S., et al. (2002) Effect of an Indo-Mediterranean diet on progression of coronary artery disease in high risk patients (Indo-Mediterranean Diet Heart Study): A randomised single-blind trial. *Lancet, 360,* 1455-1461.

Simopoulos, A.P. (2001) The Mediterranean diets: What is so special about the diet of Greece? The scientific evidence. *Journal of Nutrition, 131(Suppl.)*, 3065S-3073S.

Slattery, M. L., Boucher, K. M., Caan, B. J., Potter, J. D., & Ma, K.-N. (1998) Eating patterns and risk of colon cancer. *American Journal of Epidemiology, 148*, 4-16.

Terry, P., Hu, F. B., Hansen, H., & Wolk, A. (2001) Prospective study of major dietary patterns and colorectal cancer risk in women. *American Journal of Epidemiology, 154*, 1143-1149.

Terry, P., Suzuki, R., Hu, F. B., & Wolk, A. (2001) A prospective study of major dietary patterns and the risk of breast cancer. *Cancer Epidemiology, Biomarkers & Prevention, 10*, 1281-1285.

Trichopoulou, A., Costacou, T., Bamia, C., & Trichopoulos, D. (2003) Adherence to a Mediterranean diet and survival in a Greek population. *New England Journal of Medicine, 348*, 2599-2608.

Trichopoulou, A., Kouris-Blazos, A., Wahlqvist, M.L., Gnardellis, C., Lagiou, P., Polychronopoulos, E., et al. (1995) Diet and overall survival in elderly people. BMJ, *311*, 1457-1460.

Tseng, M., & DeVellis, R. F. (2001) Fundamental dietary patterns and their correlates among US whites. *Journal of the American Dietetic Association, 101*, 929-932.

Tucker, K. L., Chen, H., Hannan, M. T., Cupples, L. A., Wilson, P. W. F., Felson, D., et al. (2002) Bone mineral density and dietary patterns in older adults: The Framingham Osteoporosis Study. *American Journal of Clinical Nutrition, 76*, 245-252.

Tucker, K. L., Dallal, G. E., & Rush, D. (1992) Dietary patterns of elderly Boston-area residents defined by cluster analysis. *Journal of the American Dietetic Association, 92*, 1487-1491.

U.S. Department of Agriculture. (2000). *Nutrition and Your Health: Dietary Guidelines for Americans* (Home and Garden Bulletin No. 232) Fifth Edition.

Weil, A. (2000) *Eating well for optimum health*. New York: Alfred A. Knopf.

Weisburger, J. H. (1997) Dietary fat and risk of chronic disease: mechanistic insights from experimental studies. *Journal of the American Dietetic Association, 97(Suppl.)*, S16-S23.

Whichelow, M. J., & Prevost, A. T. (1996) Dietary patterns and their associations with demographic, lifestyle and health variables in a random sample of British adults. *British Journal of Nutrition, 76,* 17-30.

Wilkens, L. R. & Lee, J. (1998) Nutritional epidemiology. In P. Armitage & T. Colton (Eds.), *Encyclopedia of biostatistics* (pp. 3099-3117), Chichester, UK: Wiley.

Williams, D. E. M., Prevost, A. T., Whichelow, M. J., Cox, B. D., Day, N. E., & Wareham, N. J. (2000) A cross-sectional study of dietary patterns with glucose intolerance and other features of the metabolic syndrome. *British Journal of Nutrition, 83,* 257-266.

Willett, W. C. (2000) Nutritional epidemiology issues in chronic disease at the turn of the century. *Epidemiologic Reviews, 22,* 82-86.

Willett, W. C. (2001) *Eat, drink, and be healthy: The Harvard Medical School guide to healthy eating.* New York: Simon & Schuster.

Wirfält, E., Hedblad, B., Gullberg, B., Mattisson, I., Andrén, C., Rosander, U., et al. (2001) Food patterns and components of the metabolic syndrome in men and women: A cross-sectional study within the Malmö Diet and Cancer Cohort. *American Journal of Epidemiology, 154,* 1150-1159.

Wirfält, A. K. E., & Jeffery, R. W. (1997) Using cluster analysis to examine dietary patterns: Nutrient intakes, gender, and weight status differ across food pattern clusters. *Journal of the American Dietetic Association, 97,* 272-279.

Woo, J., Woo, K. S., Leung, S. S. F., Chook, P., Liu, B., Ip, R., et al. (2001) The Mediterranean score of dietary habits in Chinese populations in four different geographical areas. *European Journal of Clinical Nutrition, 55,* 215-220.