RELATIONSHIP OF DIETARY PATTERNS AND BONE MINERAL DENSITY AMONG OLDER U.S. MEN

by

Jaime Kalena Peter, BA, BS, MT(ASCP)

A THESIS

Presented to the Department of Public Health and Preventive Medicine

and the Oregon Health & Science University

School of Medicine

in partial fulfillment of

the requirements for the degree of

Master of Public Health

November 2005

School of Medicine
Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the MPH thesis of

Jaime Kalena Peter

has been approved

[redacted]

Jackilen Shannon, PhD

[redacted]

Lynn Marshall, ScD

[redacted]

Motomi Mori, PhD

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Acknowledgements

I would like to recognize the contribution of my thesis committee to this project. Jackie Shannon, Lynn Marshall, and Tomi Mori all contributed much in the way of professional advice, encouragement, and expertise. All three were more than happy to answer my questions and share their time with me. I would also like to thank the Osteoporotic Fractures in Men Study for allowing me access to the data, and to Dr. Katie Stone of UCSF for allowing me to alter her analysis plan and proceed with this project. Finally, I would like to thank Jolee Mougey for introducing me to the MrOS study, and Jason Walters for his constant support and encouragement.

Abstract

Osteoporosis and low bone density are important public health concerns for millions of men and women in the United States. Diet is one lifestyle factor that has been examined for prevention of low bone density and osteoporosis. Increasingly, chronic disease epidemiology uses food pattern analysis to elucidate risk factors instead of single foods or single nutrients. Reasons for this approach include accounting for intercorrelation between foods and nutrients, as well as for assessing the complexity of human diets, which consist of combinations of foods and nutrients. To determine if dietary patterns differed among community dwelling men over age 65 in 6 regions of the U.S., dietary and baseline data from the Osteoporotic Fractures in Men Study (MrOS) were examined. Of the 5,925 men included in this analysis, we found 6 distinct groups formed by factor analysis and 5 distinct groups by cluster analysis. We then used the cluster variables to examine the association between clusters and bone mineral density of the total hip. There was a marginally significant relationship between cluster and total hip BMD in a multivariate model adjusted for potential confounders ($p=0.07$). There was also a marginally significant difference in least square (LS) means between clusters 4 (red meat and fats) and 5 (heart healthy) (p for difference=0.07). Using the multivariate model constructed for the clusters, and substituting quartiles of factor scores, we found that quartiles of Factors 1 and 4 had significantly different BMD (that is, those men in the highest quartiles for the healthy factor (1) and the diet/low-fat factor (4) had significantly higher LS mean bone density than those men in the lowest quartiles (both $p<0.05$). Since this study is cross-sectional, we cannot determine if diet was truly the cause of lower or higher bone density measurements. However, the data suggest that diets higher in low-fat dairy products, fruits, vegetables, and grains may contribute to higher bone density. This information may be useful in formation of dietary guidelines for public health outreach programs for men in this age group.

Chapter 1

Introduction

Osteoporosis: Definition and Risk factors

Osteoporosis and low bone mineral density are a public health concern for millions of older U.S. men and women. According to the World Health Organization, osteoporosis is defined as "a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue with a consequent increase of bone fragility and susceptibility to fractures". The WHO's operational definition for osteoporosis is a value of bone mineral density (BMD) greater than 2.5 standard deviations below the mean for young, white adult women. Low bone density, or osteopenia, is defined as a BMD value between 1 and 2.5 standard deviations below the mean value (2004). However, it remains uncertain how to apply this criterion to men and other ethnic groups since most studies have involved white women (NIH, 2000).

Fifty five percent of the U.S. population is at risk for developing osteoporosis, including 2-3 million men. Among men, non-Hispanic whites and Asians have the highest risk for osteoporosis; it is estimated that seven percent already have the condition, while thirty five percent have low bone density. However, nineteen percent of African-American and Hispanic men are estimated to have low bone density (National Osteoporosis Foundation, 2004).

Osteoporosis is a major cause of fracture and subsequent disability. Fractures, especially at the hip, result in substantial medical expenses (WHO, 1994). In 1995,

health care expenditures attributable to osteoporotic fractures were around 13.8 billion dollars, of which 18.4% was for white men and 1.3% for nonwhite men. In 2002, the estimated cost for osteoporotic fractures was 18 billion dollars (NOF, 2004). Every year over 300,000 people are hospitalized, disabled, or require assisted care as a result of hip fractures (Fox Ray et al, 1997). Current projections estimate that the number of hip fractures will rise worldwide by nearly a factor of four between 1990 and 2050 (from 1.7 million to approximately 6.3 million) (NIH, 2000). Lifetime risk for any osteoporotic fracture in men is estimated to be between 13-22% (Johnell et al, 2004). Further, the fatality rate for hip fracture patients within one year of the injury is estimated to be 24% (Fox Ray et al, 1997). Clearly, osteoporotic fractures and low bone density are a pressing cause for concern, especially since the number of men over 70 years old is expected to double between 1993 and 2050 (Lewis, 2002). By 2030, an estimated 71.5 million people over age 65 will be living in the United States, compared to about 36 million today. This same group represented 12.4% of the population in 2000 but is expected to account for 20% by 2030 (U.S. Department of Health and Human Services, 2004).

As a modifiable trait, diet is one lifestyle factor that has been intently studied to discern risk factors for osteoporosis and low bone density (Booth et al, 2000). The majority of work has focused on the effects of calcium and vitamin D (Heaney, 1987; Cauley et al, 1999; Tucker et al, 2002). It is now accepted that calcium and vitamin D are both important contributors to bone density (Booth et al, 2000; Tucker etl al, 2002; Dawson-Hughest et al, 1991; Tucker et al, 1999). Some researchers have focused on other nutrients. Booth et al found that low intake of vitamin K was associated with increased incidence of hip fractures in a cohort of elderly men and

women (mean age 75.1 years for men) but no association with low bone mineral density (2000). Tucker et al observed that alkaline-producing dietary components, such as magnesium, potassium, and fruits and vegetables help to maintain normal BMD in elderly women and men. Since bone is a living tissue, "...it is probable that a wide spectrum of micronutrients contributes to its maintenance" (1999). In a study using data from the Framingham Heart study, Tucker et al observed that dietary patterns were associated with higher bone mineral density, and that high fruit and vegetable consumption was protective in older men (2002). Thus, overall dietary pattern may be important in determining relationships between diet, BMD, and osteoporotic fractures.

Dietary Patterns and Health

As noted by Hu, there are several potentially important limitations to the investigation of single foods or nutrients in the diet. One such limitation is that human beings eat complex diets consisting of combinations of nutrients that "are likely to be interactive or synergistic" (2002). A second limitation is that there may be high intercorrelation between nutrients that makes it difficult to assess the effects of them separately. Third, it may be easier to detect outcomes affected by multiple nutrients. As an example, Hu points to the DASH trial (Dietary Approaches to Stop Hypertension), where changing patterns of eating was more successful in lowering blood pressure than supplementation with a single nutrient (Appel et al, 1997). Fourth, there is the chance of false-positive results due to the large numbers of nutrients or foods used in typical analyses. Fifth, confounding between dietary patterns and single nutrient analysis can obscure results of studies. The potential interaction of dietary

components may not completely be adjusted for by multivariate analysis (Hu, 2002). Finally, we do not fully know the composition of foods. There may be active compounds in foods that we do not currently measure. Based on all these factors, there has been a recent shift in dietary research to not only assess individual nutrients, but the larger overall pattern of eating.

It has become clear that people eat foods, not isolated nutrients, and that they consume them in specific patterns (Schulze et al, 2003). In addition, Kant et al note that governmental recommendations for disease prevention "implicitly reflect the dietary-pattern approach by emphasizing the simultaneous change of several dietary behaviors, such as increasing fruit, vegetable, and grain intake..." (2000). This form of pattern analysis is therefore useful for determining recommendations that the public can understand and interpret (Hu et al, 2000). Analysis of dietary patterns may be able to provide a comprehensive approach to osteoporosis prevention or treatment (Hu, 2002).

Pattern analysis And Health Outcomes

In the United States, current dietary guidelines are assessed by studies that show the risk or benefit of specific dietary elements and the role of those elements in disease prevention (U.S. National Research Council, 1989). There currently is an increasing interest in the diet and health literature to examine diet as a multifaceted variable (Kant, 2004). Instead of focusing on just one specific food or nutrient, the overall pattern of diet is examined. Since 1980, at least 93 studies were published that used factor or cluster analysis to define dietary exposures; of these, 65 actually

tested hypotheses or examined associations (Newby PK et al, 2004). A variety of health conditions have been examined with this approach, including heart disease, various cancers, plasma lipid levels, overall mortality, and bone mineral density.

Prospective:

All-cause mortality: Kant et al observed that dietary patterns characterized by compliance with food-based guidelines were associated with a lower risk of all-cause mortality (2004).

Cross Sectional:

Heart disease: Quatromoni et al divided women from the Framingham Nutrition Studies into 5 specific clusters. They reported that women in the "heart healthy" cluster had the lowest probability of developing heart disease (2002).

Colon cancer: Slattery et al observed that dietary patterns were associated with risk of colon cancer; specifically, a "western" style diet was associated with the greatest risk (1998).

Gastric cancer: Kim et al observed that a "healthy" pattern derived by factor analysis was associated with a decreased risk of gastric cancer among females, while a "traditional" pattern was associated with increased risk of gastric cancer among both sexes (2004).

Plasma lipids: Newby et al reported that patterns high in reduced-fat dairy products, fruit, and fiber were inversely associated with plasma triacylglycerols with both factor analysis and cluster analysis (2004).

Bone mineral density: Tucker et al observed that high fruit and vegetable consumption in men was associated with higher BMD, while high candy consumption was associated with lower BMD in both men and women (2002).

Some studies have used interventions to show the association between food patterns and disease risk. For example, the DASH (Dietary Approaches to Stop Hypertension) study showed significant reductions in blood pressure when a low fat diet rich in fruits and vegetables was adopted (Appel et al, 1997). In addition, de Lorgeril et al observed prolonged survival and possible cancer protection in individuals assigned to a Mediterranean-style diet (1998).

Methods of Pattern Analysis

Dietary patterns cannot be directly measured in the same way as food consumption. Instead, statistical methods must be applied to food frequency questionnaire or other diet record data to characterize patterns in the cohort under study (Hu, 2002). These methods can be based on established patterns (e.g. dietary recommendations), or can be identified through data analysis. Below I will describe two methods for quantifying dietary patterns: factor analysis and cluster analysis.

Factor Analysis

The term factor analysis refers to a family of statistical techniques concerned with reducing a set of variables, often highly correlated, to a smaller number of summary variables, called factors, which characterize the most important domains represented by the set of variables. It was developed mostly in the field of psychometrics in the

social sciences for analyzing relationships between measurable entities, such as survey items. The assumptions underlying a factor model is that there are a number of unobserved variables (factors) that account for correlations among observed variables. If the latent (hidden) variables are held constant, the partial correlations among observed variables all become zero (Kachigan, 1986).

Two methods can be employed to obtain factor solutions. These methods are known as common factor analysis (CFA) and principal components analysis (PCA). The "common" in common factor analysis refers to the variance analyzed. An assumption of the method is that the variance of any single variable can be attributed to a common variance, shared by other variables in the model; a unique variance, the amount of variation associated with only a specific variable; and error variance, the variance due to unreliability in data collection, measurement error, or random error. Factors resulting from CFA are based only on common variance. Conversely, factors resulting from PCA are based on the total variance. However, both of these methods are concerned with the reduction of a large set of variables into a smaller set (Kachigan, 1986).

These two methods are widely used, and the selection of one model or the other is based both on the objectives of the analysis, and the amount of prior knowledge about the variance of the variables. PCA is more appropriate when the researcher's primary concern is prediction of the minimum number of factors needed to account for the maximum portion of variance in the original set of variables; and when prior knowledge suggests that error and unique variances represent a small proportion of

the total variance. In contrast, CFA may be more appropriate when the primary objective is simply to identify the constructs found in the original variables, and the researcher has little knowledge about the amounts of error and unique variance (Hair et al, 1998). PCA and CFA generally result in similar solutions if the number of variables exceeds 30 (Hair et al, 1998). Although mathematically different, the two methods often produce comparable results (Kachigan, 1986)

Once factors are derived, regardless of the methodology used, further data reduction can be achieved by calculation of factor scores. Factor scores can quantify individual cases on a continuum using a z-score scale (mean 0, standard deviation 1). These scores can be substituted for the original variables in further analyses (Kachigan, 1986).

Cluster Analysis

Cluster analysis is a second method that can be used to quantify patterns in a set of variables. It is an exploratory data analysis tool that classifies objects so that each object is very similar to others in the same cluster with respect to some predetermined characteristic (Hair et al, 1998). Thus, individuals are placed into non-overlapping groups on the basis of a similar characteristic (Kant, 2004). With cluster analysis, the objective is "joining together different subjects to subgroups" (Koffmann et al, 2000). Patterns are identified by grouping individuals with similar characteristics, which produces homogeneous and non-overlapping exposure categories (Wirfalt et al, 1999). Each subject will belong to only one cluster at the end of the procedure.

In summary, factor analysis aggregates foods into groups based on correlations between foods; individuals receive a score for all derived factors. In contrast, cluster analysis separates people into mutually exclusive groups based on differences in mean food intakes, or clusters (Newby et al, 2004). Both methods are considered to be "*a posteriori*" (knowledge based on experience) because the eating patterns are derived from modeling data already obtained. Patterns identified with each method do not necessarily represent an ideal eating pattern (Hu, 2002). However, these techniques may provide an alternative approach to the assessment of single foods or nutrients in epidemiologic studies (Quatromoni et al, 2002).

The objectives of this study were 1) to use factor and cluster analysis to determine whether different dietary patterns exist among MrOS participants who completed the baseline food frequency questionnaire, and 2) to perform a cross sectional analysis to investigate if dietary patterns derived from the baseline dietary data are related to bone mineral density of the total hip. We hypothesized that different dietary patterns did exist in this study population, and that those differences might be associated with differences in bone mineral density.

Chapter 2

Methods

Participants

The Osteoporotic Fractures in Men (MrOS) study is a multi-center, prospective cohort study. The primary goal of the MrOS study is to quantify risk factors for fractures in U.S. men ages 65 years and older. Secondary aims include examination of risk factors for other age-related conditions such as prostate cancer. Community-dwelling men were enrolled from 6 diverse cities (Palo Alto and San Diego CA; Birmingham AL; Pittsburgh PA; Minneapolis MN; and Portland OR). Enrollment was completed between March 2000-2002. Around 1000 men were enrolled from each site for a total of 5,995 participants. The study progress and data collection are managed by the Coordinating Center at the California Pacific Medical Center/University of California, San Francisco, and the Administrative Center at Oregon Health and Science University. The MrOS Steering Committee is responsible for the scientific direction of the study (Orwoll E et al, 2005).

Each participating clinical site had their own recruitment strategies to enroll men into the cohort, with an emphasis on recruitment of minority groups. Recruitment strategies included mailings to age-eligible residents identified from DMV and voter registration lists; community and senior newspaper advertisements; word of mouth; and targeted presentations. Criteria for inclusion into the study consisted of ability to walk without assistance from another person, absence of bilateral hip replacements, ability to provide self-reported data, likelihood of completing the study (i.e. residence

near a clinical site for the study's duration), absence of a medical condition that would result in imminent death, and ability to understand and sign an informed consent document (Blank JB et al, 2005).

The Institutional Review Board at each of the clinic sites approved the MrOS study. In addition, written informed consent was obtained from each study participant.

Dietary Assessment

To have the most accurate and reliable nutritional data, a standardized food frequency questionnaire (FFQ) was utilized. Block dietary questionnaires (Block Dietary Data Systems, Berkeley, CA) have been validated and are in use by over 700 registered research and public health groups (Berkeley Nutrition Services, ND). The version of the FFQ used in MrOS was designed specifically to capture the most frequently consumed sources of calcium and other key nutrients in men over age 65 in the past year. Participants completed the 67-item food frequency questionnaire at baseline. There were two types of questions for each food item—the frequency ("how often") and the portion ("how much"). Some of the serving size options were illustrated with quantities (e.g. ½ plate) labeled A, B, C, and D. There were also modifying questions present, such as "how often do you eat chicken without skin" or "how often do you eat meat trimmed of fat?" Each FFQ was sent to Berkeley Nutrition Services for nutrient analysis. Nutrient and caloric intake values were returned, as well as number of servings of foods based on USDA Department of Agriculture food pyramid guidelines. A data set consisting of grams of food was also returned, as well as information on supplement use.

Important measurements from the dietary data include daily calories (measured in kcals), calcium (mg/day), and vitamin D (IU/day). Use of calcium and/or vitamin D supplements was recorded as well; these were added to dietary levels for an overall measure of total calcium and vitamin D intake. We used these FFQ data on foods consumed for the subsequent factor and cluster analyses.

For the current analysis, participants were excluded if they had implausible daily caloric intake (less than 400 kCal/day, greater than 6000 kCal/day). This led to the exclusion of 70 participants, leaving 5,925 men.

Measurement of Bone Mineral Density

Bone mineral density (BMD) was assessed with Dual Energy X-ray Absorptiometry (DEXA) technology with machines of the same make and model at each site (QDR 4500W, Hologic Inc., Bedford, MA). BMD was measured at the lumbar spine (L1-L4) as well as total hip and subregions (trochanter, femoral neck). Standardized procedures were used at each site to assure proper participant position and scan analyses. Furthermore, a random sample of all scans was reviewed by densitometry technicians at the data-coordinating center. The precision of the DEXA machines was measured prior to the baseline visit; cross calibration studies found that the maximum percent difference in total mean spine BMD was 1.4% (Cauley et al, 2005). All measurements are reported in grams per square centimeter (g/cm^2).

Other Baseline Measurements

Each study participant attended a comprehensive baseline clinic visit. This visit included collection of serum, urine, and DNA samples as well as assessment of nutritional, lifestyle, and activity level. A self-administered questionnaire was used for lifestyle, medical, and physical activity history, while a Block food frequency questionnaire was used to obtain nutritional information. Assessment of vertebral fractures by x-rays of the spine; recording of height, weight, grip strength, and leg strength; and administration of tests for visual, neuromuscular, and mental function were also completed.

Race/Ethnicity was self-reported; categories included Caucasian; Black/African-American; Asian; Hispanic or Latino; Native Hawaiian/Pacific Islander; American Indian/Alaskan Native; and multiracial. For this analysis, the first 4 groups were used and the 5^{th}, due to small numbers, consisted of "other" (the last 3 categories). Weight was measured for each participant without shoes using a balance beam scale, while height was measured with a Harpenden stadiometer (DyFed, UK). Body mass index was calculated using both of these measurements, and has units of kg/cm^2. Height and weight loss since age 25 was measured by subtracting self-reported weight and height at 25 years from current weight and height.

Alcoholic beverage consumption (assessed via questionnaire) was measured as the average number of drinks per week, while smoking status was lifetime history (current, past, or former smoker). Physical activity in the past seven days was

measured with the PASE scale (Physical Activity Scale for the Elderly); this variable is continuous and is designed to capture leisure activities such as gardening. Assessment of whether subjects walk for exercise was completed by asking the question "Do you take walks for exercise, daily or almost every day" (yes/no)?

An extensive medical history was obtained from participants at baseline. All prescription and non-prescription medicines were brought to the clinic for verification. The use of a study-specific medication dictionary enabled categorization of medications (whether generic or name brand) from the product containers. When assessing if use was current, 30 days was used for the cutoff. However, dose and duration were not obtained; the exception was to elicit whether the participant had ever taken medication for osteoporosis.

MrOS participants were asked whether they had been told by a health provider that they suffered from one of the following medical conditions: angina, arthritis, hypertension, cancer, cataracts, CHF, COPD, diabetes, dizziness, glaucoma, osteoarthritis, thyroid problems (high or low), kidney stones, heart attack, osteoporosis, Parkinson's disease, prostatitis, gastrectomy, or stroke. Fracture history of each subject was also obtained (type of bones broken, fracture history after age 50, traumatic vs. nontraumatic fractures). Maternal and paternal history of hip fracture was obtained. Participants were also asked about self-reported health status. This was answered in terms of good/excellent vs. fair, poor, very poor.

Statistical analysis

This section will describe more specifically the steps followed to obtain factor solutions and cluster groups, respectively, which were described generally in the Introduction section, pp 6-8.

Factor Analysis

Grams of foods consumed from the FFQ were converted to daily servings using standard gram weights for food items. Daily servings of individual foods were used as the variables in the analysis. The PROC FACTOR procedure in SAS (v 9.1, SAS Institute, Cary, NC) was utilized for the analyses. The selection of the model options are described below, as these are the main determinants of how many factors were extracted. Common factor analysis was performed to obtain factor solutions.

Latent Root Criterion: The rationale for this most commonly used option is that any individual factor should account for the variance of at least one variable if it is to be retained. Each variable contributes a value of 1 to the total eigenvalue, so only factors having latent roots (eigenvalues) greater than 1 are retained (Kachigan, 1986). **Figure 1** provides a visual representation of this idea; the solid line indicates that with this option, 6 factors are retained.

Scree test: This test is derived by plotting eigenvalues against the number of factors in their order of extraction; the shape of the curve is used to evaluate the cutoff point. The point at which the curve straightens out is thought to indicate the maximum number of factors to extract (University of Texas at Austin Statistical Services, 1995). In **Figure 1**, there is not a very clear point at which the curve bends; it would be somewhere around 15 factors, as demonstrated by the dashed line. This was felt to be too great for interpretability; therefore this test was not utilized.

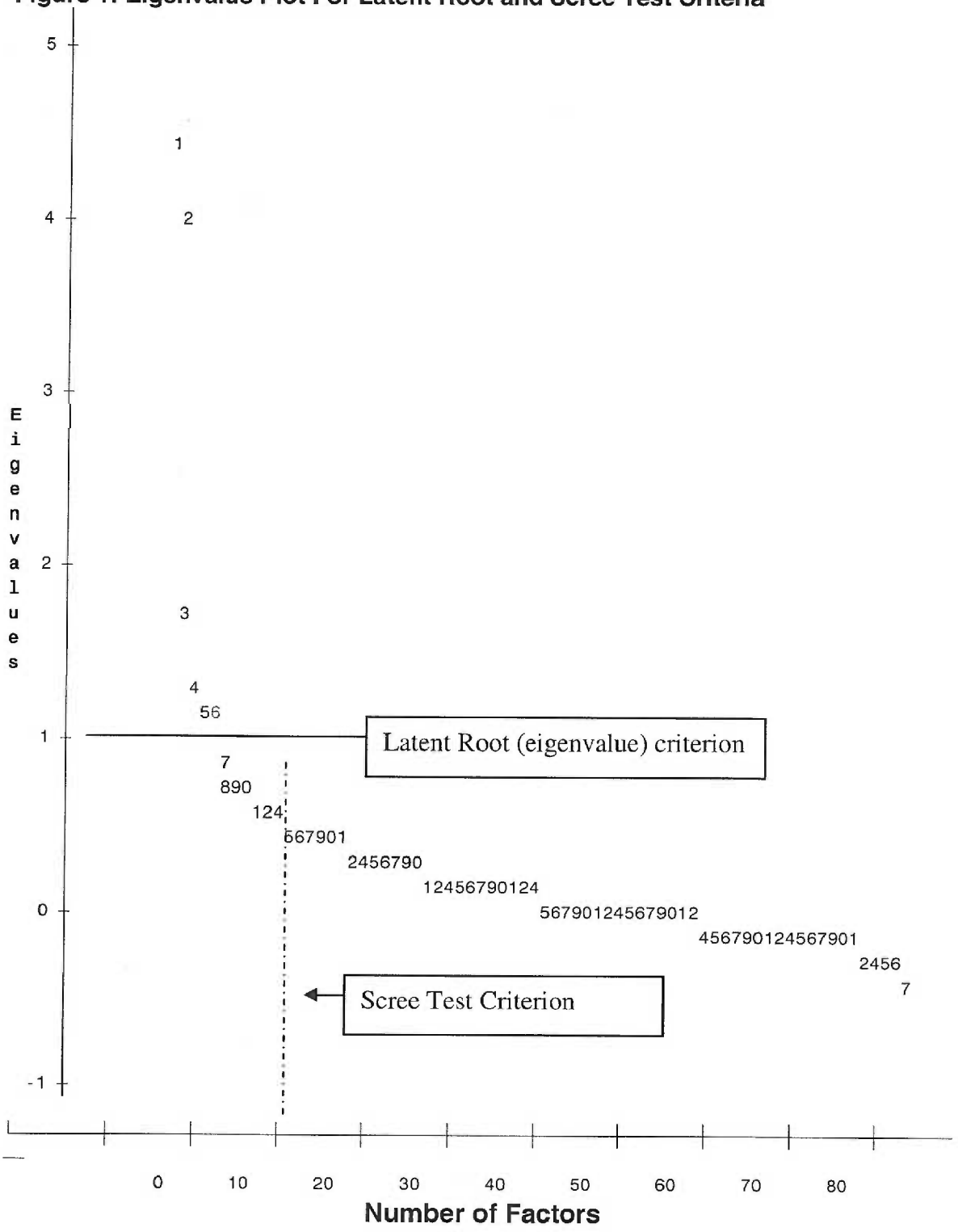
Interpretability: The factor solutions should be evaluated for meaningfulness. This is a subjective decision on part of the investigator (University of Texas at Austin Statistical Services, 1995)

Rotation of factors: The initial factor pattern matrix is not unique, and the same matrix can be produced many ways by rotating the reference axes of the factor solution. An orthogonal rotation (where the angle between the reference axes of factors are maintained at 90 degrees) is the simplest case (University of Texas at Austin Statistical Services, 1995). This option was used in the current analysis (varimax option in SAS).

Factor loadings: This is an *n* by *m* matrix of correlations, where *n* is the number of food variables and *m* is the number of retained factors. The meanings for the factors are inferred from the variables that significantly load on the rotated factors. A rule of thumb is that factor loadings of greater than 0.3 (in absolute value) are considered significant (University of Texas at Austin Statistical Services, 1995). Factor loadings greater than 0.2 were used in the current analysis.

The PROC FACTOR procedure was used with the specifications for eigenvalues >1 and orthogonal (varimax) rotation. Based on these results, a 6-factor solution was selected. All significant loadings were identified, and factors were named based on the foods that loaded most positively on the factor. In addition, factor scores were calculated for each subject for each of the 6 factors; the standardized intakes of all the foods were weighted by factor loadings and summed, then standardized (mean 0, SD 1) using the PROC SCORE procedure. These factor scores were then divided into quartiles for further analysis.

Figure 1: Eigenvalue Plot For Latent Root and Scree Test Criteria



Cluster analysis

In contrast to factor analysis, cluster analysis is used to place individuals into groups on the basis of a similar characteristic (food intake). Whereas factor analysis aggregates specific food items, cluster analysis aggregates individual persons into subgroups (clusters) with similar dietary patterns.

The procedure PROC FASTCLUS in SAS (v9.1, SAS Institute, Cary, NC) was used for this analysis. This procedure uses a K-means method to classify participants into predetermined numbers of clusters, based on Euclidian distances. The "K" in K-means is the number of clusters that are preselected by the investigator. Thus, this method produces k clusters of maximum distinction. Euclidian distance refers to the geometric distance in a multidimensional space (i.e., square root of the sum of the differences in each variable). In cluster analysis, the distances are the similarities or differences in the variables that are to be clustered (Garson, 2004).

The first step by the software is to scan the dataset for initial cluster seeds. These are selected as a first "guess" of the means of each cluster. Temporary clusters are then formed by assigning each participant to the nearest seed. This algorithm is repeated multiple times, until the final cluster seeds are equal to the cluster means. Thus, participants are moved between clusters while the program runs; new cluster means are calculated until "the distances between observations within a cluster are smaller than the distances between cluster means" (Wirfalt et al, 1997). Cluster analysis is sensitive to outlying values, and men with implausible daily caloric intake, as described above, were excluded.

As mentioned above, the number of clusters must be preselected in this type of analysis. Since no *apriori* information was available about the actual clusters that might be present in the data, a number of steps were taken to find an appropriate number. Many runs of the clustering procedure were undertaken, with number of clusters specified from 20 to 2. Ratios of between cluster variance to within cluster variance [Overall $R^2/(1-R^2)$] were examined for each cluster run. A plot was constructed of these values versus the number of clusters in that run. Similar to a scree plot, the point where the next higher number of clusters would not give substantially better separation was chosen. Finally, discriminant analysis plots were made with different numbers of cluster solutions. The PROC CANDISC (canonical discriminant analysis) procedure was used to generate the data. With this procedure, the output data set from each cluster run was used to find maximal separation of the groups. Each cluster solution was the grouping variable, while the food groups were the quantitative variables. The procedure derives canonical variables, which are linear combinations of the food group variables that summarize between-class variation. Plotting pairs of canonical variables can aid in visual interpretation of group differences (SAS Institute, 2002). Finally, clusters were examined for nutritionally meaningful separation. Using all of these criteria, a 5-cluster solution was selected.

Statistical Analysis: Assessing Association with Bone Density

Descriptive Comparisons

Quartiles of factor scores (highest vs. lowest) and clusters were separately tested for significant differences across variables of interest. Descriptive comparisons across

quartiles of factor scores and clusters were obtained by regressing continuous demographic and lifestyle measurements in the general linear models procedure (adjusted for age). For categorical demographic and lifestyle variables, Pearson's chi-square test was utilized. Significance level was set at a p-value of ≤0.05.

Covariate/Confounder Assessment

Variables observed in previous analyses (Cauley et al, 2005) that were associated with hip bone mineral density were assessed in age-adjusted univariate models using the general linear models procedure (Proc GLM). Because of the large number of potential confounding variables, significance level was set at an alpha of 0.05. Those variables were then examined in relation to the cluster variables using chi-squares and one-way ANOVA (for categorical/categorical, and continuous/categorical, respectively). Variables were considered to be potential confounders if they were associated with total hip BMD and quartiles of factor scores, or total hip BMD and cluster variables.

Multivariable Modeling

The multivariate model was constructed manually in two stages using linear regression. First, 7 subgroups of potential variables to include in the multivariate model were independently analyzed. These subgroups included demographics; body composition; medical history (including prevalent medical conditions); medication use; dietary variables; activities of daily living; and fracture history. These subgroups helped to manage the large number of potential variables to include, and to limit problems with correlated variables. Within each subgroup, variables were ranked

according to an age-adjusted association with total hip bone density. The cluster variable was included in each model because it was the primary exposure measure. A "best" model was chosen by examining the change in the model R square value (with increasing values ranking higher), as well as by examining p values for each variable added in the model (with significance at $p \leq 0.05$).

Once the "best" model within each subgroup was determined, the models were combined by sequential addition to the largest subgroup (medication use). With the addition of each group, significance was assessed using the criteria described in the above paragraph. When all variables from the best subgroup models were added, any variables not meeting an inclusion criterion of $p \leq 0.05$ were removed from the combined model (beginning with those with the highest p-value). Interaction terms were assessed by using biologically plausible criteria, as well as suspected possible dietary variable interactions. A p value of ≤ 0.05 was used for significance. To ensure that no variable was erroneously taken out, variables removed were added back to the final model and reassessed.

Results for models with total hip bone density as the dependent variable are presented as least squares means, adjusted for multiple comparisons by the Tukey-Kramer option in the general linear models procedure.

When a final model was selected, quartiles of factor variables (highest vs. lowest) were used individually to replace the cluster variables. This resulted in an additional 6 models. This approach was chosen for two reasons: first, to limit the amount of modeling necessary for this project; second, to see how our ability to predict or

explain variance in bone density changed when using this different approach to finding dietary patterns.

Residual analysis was performed to check for outliers in the data. Dummy variables were created manually from the cluster variables; clusters 1-4 were then used in a linear regression analysis with cluster 5 as the reference. Quantile-Quantile (QQ) plots were examined for outlying data points. These three outliers were then removed from the dataset, and modeling was performed to see how parameter estimates changed, if at all.

All analyses were conducted using SAS version 9.1 (SAS Institute, Cary, NC).

Chapter 3

Results

Table 1 presents the baseline characteristics of the 5,925 participants who comprised the population for the present study. As shown in Table 1, MrOS participants were relatively heavy (mean body mass index 27.4); in addition, white race predominated (89.6%). Only a small percentage of men were current smokers (3.4%); average alcohol intake per weeks was just over 4 drinks. A high percentage of MrOS participants were currently married (82.4%), and a similar percentage rated their health as good or excellent (85.8%).

Table 1. Characteristics of the MrOS cohort used in current analyses at the time of the bone mineral density measurement[1]

Age (yr)	73.6±5.8[2]
BMI (kg/m^2)	27.4±3.8
Race (%)	
White	89.6
African American	4.0
Asian	3.0
Hispanic	2.1
Other	1.2
Smoking Status (%)	
Past smoker	59.2
Current	3.4
Health Status (%)	
Good/excellent	85.8
Fair/poor/very poor	14.2
Daily Caloric Intake (kcal)	1619.5±630.3
Average drinks/week (#)	4.3±6.8
Calcium supplement user (%)	33.9
Vitamin D supplement user (%)	12.0
Calcium intake (mg)	1141.5±589.6
Vitamin D intake (IU)	391.7±245.4
Marital Status (%)	
Married	82.4
Divorced	5.2
Widowed	8.8
Bone mineral density (g/cm^2)	
Total hip	0.96±0.14
Femoral Neck	0.78±0.13
Trochanteric	0.76±0.13

[1] Sample size varied from 5918 to 5925 due to missing values
[2] Mean ± SD; other values presented as percentages

Factor loadings for the 6 food factors found in this analysis, along with the names assigned to each factor, are shown in **Table 2 (pp 27)**. The factor loadings are interpreted like a correlation coefficient, where the most positive value contributes the most to a factor score, and the most negative value contributes the least to the factor score. The food patterns derived from the analyses are as follows: "healthy," in which fruits, vegetables, and legumes loaded the highest; "sweets/convenience," where baked goods and fried foods loaded the highest; "meat/western" where meats and cheeses loaded the highest; "diet/low fat" where diet salad dressings, low fat meats, and nonfat milk loaded the highest; "fatty meat" where meats and untrimmed meats loaded the highest; and "Mediterranean" where salad, tomatoes, spaghetti, cheese, and fish loaded the highest. Factor one was the predominant food pattern in this study population; it explained 3.4% of the variance in intake. Factors two and three each explained over 2% of the intake (2.8% and 2.4% respectively), while factors four through six together explained just over 5% of the variance in intake (data not shown).

Baseline sample characteristics by quartiles of each factor score are presented in **Tables 3 and 4 (pp 28-29)**. Results are presented for the lowest (Q1) and highest (Q4) quartile of each food pattern. Continuous variables are expressed as mean ± standard deviation, while categorical variables are presented as percentages. **Table 5 (pp 30)** presents p-values for overall association of quartiles of factor scores and selected categorical variable categories.

Continuous variables

There was little variation in selected demographic, lifestyle, and medical variables examined by quartile of factor score. Higher mean daily calories were consumed by the men in the 4th quartiles of factors two and three. Body mass index was similar across all factors, but again men in the 4th quartile of factors two and three had the highest BMI (27.6). Men in the 4th quartile of factor one had the highest daily intake of calcium (1174 mg); they also had the highest daily intake of vitamin D (399 IU). Men in all factors had very similar Teng mental scores, ages, heights, and PASE scores.

Categorical variables

Among demographic and lifestyle variables, age groups significantly differed across quartiles of factor scores (except for quartiles of factor one, where p=0.06; quartiles of factor two, where p=0.13; and quartiles of factor four, where p=0.61); site, education; race; marital status; and smoking status (all p<0.05). See Table 5, below.

Variation in medication and medical history variables across factor scores is greater than demographic and lifestyle variables, as can be ascertained from Table 5. History of osteoarthritis (hip, hand, knee) was significantly different across quartiles of factor one (p=0.04) but not any other factors. Conversely, history of heart disease was significantly different across quartiles of factor four (p=0.0003) but not any other factors. History of central nervous medication use was significantly different across quartiles of factor two (p=0.03) and quartiles of factor six (p=<0.01). Use of oral

hypoglycemic agents was significantly different across quartiles of factors one, three, four, and five.

In the activities of daily living variables, ability to walk 2-3 blocks was significantly different across quartiles of factors two and three only ($p<0.01$); ability to prepare one's own meals significantly different across quartiles of factor three only ($p=0.02$).

Table 2

Factor loadings for 6 food patterns identified at baseline in the MrOS study

Food [1]	Factor loading[2]
Factor 1: "Healthy"	
Carrots	0.53
Broccoli	0.51
Spinach	0.51
Other Vegetables	0.46
Green beans	0.46
Baked/pinto beans	0.43
Coleslaw/cabbage	0.42
Tofu	0.39
Other fresh fruit	0.38
Corn	0.37
Apples/pears	0.34
Raw tomatoes	0.32
Rice	0.32
Fish	0.32
Sweet potatoes	0.31
Oranges	0.30
Salad	0.29
Vegetable soup	0.28
Hot cereals	0.25
Bananas	0.25
Yogurt/frozen yogurt	0.20
Factor 2: "Sweets/convenience"	
Doughnuts/pastries	0.47
Cookies	0.46
Chocolate candy	0.41
Ice cream	0.38
Pies	0.38
Crackers	0.32
French fries	0.31
Cheese dishes	0.32
Hamburger	0.30
Margarine	0.30
Ham and lunch meats	0.29
Potatoes not fried	0.29
Salty snacks	0.28
Gravy	0.28
Pizza	0.28
Peanut butter	0.28
Mayonnaise	0.27
White bread	0.26
Butter	0.24
Hot dogs	0.22
Spaghetti	0.21
Factor 3: "Meat/Western	
Bacon and sausage	0.59
Eggs	0.55
French fries	0.37
Hot dogs	0.36
Hamburger	0.35
Gravy	0.29
Pork, fat-trimmed	0.28
Beef, fat-trimmed	0.25
Fried chicken	0.24
Cheese dishes	0.23
Nonfat milk	-0.21
Factor 4: "Diet/Low fat"	
Diet salad dressings	0.45
Low fat lunch meats	0.40
Low fat cheese	0.35
Low fat ice cream	0.32
Nonfried chicken, no skin	0.32
Low fat cookies	0.31
Nonfat milk	0.29
Dark bread	0.22
Ice cream	0.20
Cheese	-0.22
Butter	-0.23
Salad dressing	-0.27
Factor 5: "Fatty meat"	
Beef	0.69
Pork	0.67
Fried Chicken	0.32
Chicken not fried	0.24
Nonfried chicken, no skin	0.21
Beef, fat-trimmed	-0.47
Pork, fat-trimmed	-0.48
Factor 6: "Mediterranean"	
Salad	0.55
Salad dressing	0.52
Raw tomatoes	0.40
Spaghetti	0.26
Cheese	0.25
Fish	0.23
Shellfish	0.22
Beef, fat-trimmed	0.21

[1] Foods with absolute values <0.2 are omitted for simplicity

[2] Correlation coefficients; factor loadings represent the magnitude and direction of association with factors and range from –1 to +1.

Table 3. Selected sample characteristics at baseline for the lowest and highest quartiles of each food pattern for 5,925 men in the MrOS study (Factors 1-3)*

Sample Characteristic**	Factor 1 "Healthy"		Factor 2 "Sweets"		Factor 3 "Western"	
	Q1	Q4	Q1	Q4	Q1	Q4
Age (yr)	73.5±5.8	73.7±5.8	73.7±5.8	73.8±6.0	74.0±5.8	73.4±5.9
Height (cm)	174.0±6.7	174.5±6.8	174.2±6.8	173.9±6.8	174.1±6.8	174.0±6.9
Weight (kg)	83.3±13.3	82.9±13.1	83.0±13.1	83.4±13.0	82.6±13.0	83.7±13.5
Daily Calories (kcal)	1604±646	1648±637	1588±678	1685±669	1599±631	1670±678
Daily Calcium (mg)	1121±594	1174±595	1146±616	1128±572	1160±615	1130±582
Daily Vitamin D (IU)	387±244	399±249	397±256	387±244	396±248	388±245
PASE score	144.8±68.0	146.8±69.1	146.0±70.5	146.6±68.8	145.8±68.5	147.4±69.1
Drinks/week (#)	4.1±6.7	4.5±7.0	4.5±7.2	3.9±6.5	4.4±6.5	4.0±6.8
Body mass index (kg/m^2)	27.4±3.8	27.2±3.7	27.3±3.7	27.6±3.82	27.2±3.6	27.6±3.9
Teng Mental Score	93.4±5.9	93.4±5.7	93.5±5.7	93.2±5.8	93.6±5.5	93.0±6.3
Ethnicity [n(%)]						
White	1338(90.4)	1313(88.6)	1288(87.0)	1353(91.3)	1321(89.2)	1344(90.7)
African American	59(4.0)	62(4.2)	69(4.7)	49(3.3)	57(3.8)	63(4.2)
Asian	35(2.4)	53(3.6)	65(4.4)	32(2.2)	44(3.0)	36(2.4)
Hispanic	29(2.0)	39(2.6)	40(2.7)	34(2.3)	44(3.0)	18(1.2)
Other	19(1.3)	15(1.0)	19(1.3)	14(<1)	15(1.0)	21(1.4)
Site [n(%)]						
Birmingham	212(14.3)	268(18.1)	206(13.9)	270(18.2)	149(10.1)	390(26.3)
Minneapolis	334(22.6)	183(12.3)	159(10.7)	313(21.1)	241(16.3)	274(18.5)
Palo Alto	185(12.5)	304(20.5)	350(23.6)	153(10.3)	320(21.6)	149(10.0)
Pittsburgh	305(20.6)	209(14.1)	145(9.8)	380(25.6)	219(14.8)	286(19.3)
Portland	222(15.0)	259(17.5)	302(20.4)	206(13.9)	272(18.4)	191(12.9)
San Diego	222(15.0)	259(17.5)	319(21.5)	160(10.8)	280(18.9)	192(13.0)
Education [n(%)]						
<High School	93(6.3)	95(6.4)	93(6.3)	117(7.9)	93(6.3)	99(6.7)
HS, some College	649(43.8)	572(38.6)	541(36.5)	634(42.8)	538(36.3)	676(45.6)
Completed college	268(18.1)	275(18.6)	282(19.0)	254(17.1)	281(19.0)	256(17.3)
>College	470(31.8)	540(36.4)	565(38.1)	477(32.2)	569(38.4)	451(30.4)
Marital Status [n(%)]						
Currently Married	1217(82.2)	1222(82.5)	1198(80.9)	1224(82.6)	1221(82.4)	1238(83.6)
Other	263(17.8)	260(17.5)	283(19.1)	258(17.4)	260(17.6)	293(19.8)
Smoking status [n(%)]						
No	550(37.2)	549(37.0)	580(39.2)	539(36.4)	588(39.7)	529(35.7)
Past	874(59.0)	880(59.4)	854(57.7)	887(59.9)	846(57.1)	901(60.8)
Current	56(3.8)	53(3.6)	47(3.2)	56(3.8)	47(3.2)	52(3.5)

***Factors are standardized continuous variables; each person has a score for each factor**

****Continuous variables are mean±sd; categorical variables are n(%)**

Table 4. Selected sample characteristics at baseline for the lowest and highest quartiles of each food pattern for 5,925 men in the MrOS study (Factors 4-6)*

Sample Characteristic**	Factor 4 "Diet/lowfat"		Factor 5 "Fatty meat"		Factor 6 "Mediterranean"	
	Q1	Q4	Q1	Q4	Q1	Q4
Age (yr)	73.7±5.8	73.7±59	73.7±6.0	73.6±5.9	73.6±5.7	73.9±6.1
Height (cm)	174.0±6.9	174.4±6.8	174.1±7.0	174.1±6.7	174.1±6.8	174.1±6.8
Weight (kg)	82.6±13.2	83.2±12.9	83.5±13.6	83.6±13.1	83.5±13.5	82.4±12.6
Daily Calories (kcal)	1641±676	1608±638	1630±602	1639±652	1645±686	1617±632
Daily Calcium (mg)	1160±592	1138±582	1133±585	1162±585	1110±589	1152±590
Daily Vitamin D (IU)	393±243	392±246	395±246	391±241	383±250	396±247
PASE score	146.6±69.4	147.2±67.9	146.2±66.0	146.0±69.0	146.2±69.8	146.0±67.0
Drinks/week (#)	4.1±6.4	4.4±7.6	4.4±6.9	4.1±6.6	3.9±6.6	4.40±6.4
Body mass index (kg/m^2)	27.2±3.8	27.3±3.7	27.4±4.0	27.5±3.9	27.5±39	27.2±3.7
Teng Mental Score	93.2±5.8	93.6±5.3	93.2±6.3	93.2±6.0	93.1±5.9	93.2±6.1
Ethnicity [n(%)]						
White	1305(88.2)	1350(91.1)	1355(91.4)	1304(88.0)	1320(89.1)	1340(90.4)
African American	66(4.5)	52(3.5)	56(3.8)	65(4.4)	775.2)	45(3.0)
Asian	49(3.3)	34(2.3)	32(2.2)	63(4.2)	40(2.7)	44(3.0)
Hispanic	40(2.7)	26(1.8)	20(1.3)	32(2.2)	23(1.5)	36(2.4)
Other	20(1.3)	19(1.3)	19(1.3)	17(1.1)	21(1.4)	17(1.1)
Site [n(%)]						
Birmingham	250(16.9)	207(14.0)	243(16.9)	264(17.8)	424(28.6)	118(18.0)
Minneapolis	184(12.4)	309(20.9)	291(19.6)	274(18.5)	238(16.1)	197(13.3)
Palo Alto	276(18.6)	216(14.6)	216(14.6)	269(18.2)	177(11.9)	3239(21.8)
Pittsburgh	266(18.0)	269(18.2)	315(21.3)	211(14.2)	233(15.7)	292(19.7)
Portland	276(18.6)	215(14.5)	226(15.2)	254(17.1)	243(16.4)	242(16.3)
San Diego	228(15.4)	265(17.9)	191(12.9)	209(14.1)	166(11.2)	310 (20.9)
Education [n(%)]						
<High School	112(7.6)	88(5.9)	97(6.5)	84(5.7)	126(8.5)	94(6.3)
HS, some College	599(40.5)	615(41.5)	647(43.7)	581(39.2)	615(41.5)	559(37.7)
Completed college	266(18.0)	263(17.8)	253(17.1)	269(18.2)	253(17.1)	291(19.6)
>College	503(34.0)	515(34.8)	485(32.7)	547(36.9)	487(32.9)	538(36.3)
Marital Status [n(%)]						
Currently Married	1238(83.6)	1202(81.2)	1213(81.8)	1209(81.6)	1212(81.8)	1221(82.4)
Other	242(16.3)	279(18.8)	269(18.1)	272(18.4)	269(18.2)	261(17.6)
Smoking status						
[n(%)]	552(37.3)	563(38.0)	530(35.8	518(35.0)	538(36.3)	593(40.0)
No	874(59.0)	863(58.3)	899(60.7)	924(62.4)	889(60.1)	843(56.9)
Past	54(3.6)	54(3.6)	53(3.6)	38(2.6)	53(3.6)	46 (3.1)
Current						

*** Factors are standardized continuous variables; each person has a score for each factor**
****Continuous variables are mean±sd; categorical variables are n(%)**

Table 5. Results of overall factor quartiles (F1Q...F6Q) and selected categorical variables

Variable Category	p value* F1Q	F2Q	F3Q	F4Q	F5Q	F6Q
Demographics/Lifestyle						
Age group	0.06	0.13	<.01	0.61	<.01	<.01
Site	<.01	<.01	<.01	<.01	<.01	<.01
Education	<.01	<.01	<.01	<.01	<.01	<.01
Race	<.01	<.01	<.01	<.01	<.01	<.01
Marital Status	.0005	0.03	<.01	.0006	<.01	<.01
Smoking status	<.01	0.005	<.01	<.01	<.01	0.07
Medical history						
Osteoarthritis	0.04	0.20	0.92	0.16	0.78	0.38
Heart Disease	0.13	0.24	0.46	0.0003	0.28	0.76
Medication use						
Central nervous system†	0.71	0.03	0.36	0.18	0.51	<.01
Corticosteroids	0.32	0.21	0.62	0.89	0.74	0.50
Osteoporosis medication	0.41	0.46	0.62	0.04	0.16	0.60
Ace inhibitors	0.41	0.24	0.05	0.14	0.98	0.63
Beta blockers	0.50	0.54	0.32	0.02	0.23	0.59
Oral hypoglycemics	0.003	0.39	<.01	0.0001	0.004	0.67
Loop diuretics	0.77	0.61	0.36	0.14	0.23	0.002
Proton pump inhibitors	0.18	0.04	0.68	0.97	0.44	0.002
Statins	0.004	0.001	0.02	<.01	0.03	0.16
Activities of daily living						
Walk 2-3 blocks	0.032	0.0002	<.01	0.77	0.92	0.09
Prepare own meals	0.73	0.44	0.02	0.77	0.92	0.09

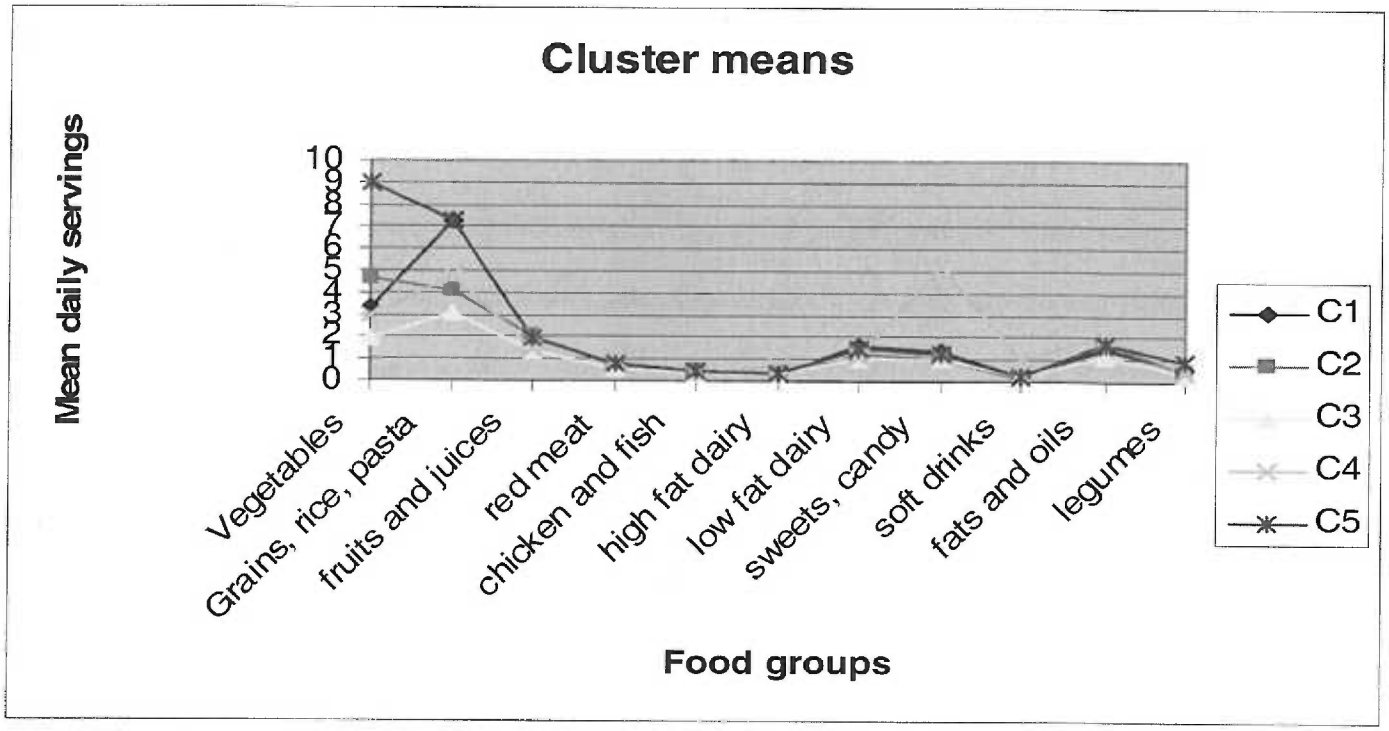
*from Pearson chi-square analysis
†Selective serotonin reuptake inhibitors, benzodiazepines, or non-benzodiazepine anticonvulsant drugs

The results of the cluster analysis are presented in **Table 6**. A solution of five clusters was chosen using the previously described criteria. The clusters are named according to the food or type of food that contributed the highest number of servings. These clusters consist of bread & pasta (1), light eating (2), many foods (3), red meat and fats (4), and heart healthy (5). Cluster 3 had the most subjects (2848), followed by Cluster 2 and Cluster 1 (1364 and 1006 subjects, respectively). Clusters 4 and 5 had the least number of subjects (351 and 356, respectively). A graphical representation of the cluster means for the different food groups is presented in Figure 2, below.

Table 6
5 clusters of participants according to food consumption showing number of men per cluster, food groups per cluster, and average daily servings

	Average Servings (mean ± sd)				
Cluster Grouping	**C1 (n=1006) Bread/pasta**	**C2 (n=1364) Light eating**	**C3 (n=2848) Many foods**	**C4 (n=351) Red meat & fats**	**C5 (n=356) Heart healthy**
Food Groups					
Vegetables	3.27±1.36	**4.69±1.12**	1.89±0.85	2.94±1.46	**8.96±2.08**
Grains, rice, pasta	**7.29±1.64**	**4.11±1.14**	3.11±1.17	4.76±1.89	**7.29±2.33**
Fruits & juices	1.92±0.96	**2.03±0.97**	1.42±0.84	1.49±0.87	1.92±1.03
Red meat	0.74±0.55	0.7±0.53	0.52±0.40	**0.99±0.72**	0.74±0.78
Chicken & fish	0.45±0.36	0.48±0.38	0.29±0.26	0.39±0.40	0.45±0.65
High fat dairy	0.38±0.52	0.33±0.46	0.31±0.41	0.73±0.81	0.38±0.51
Low fat dairy	**1.54±1.12**	1.3±1.05	1.06±0.94	1.09±1.14	1.48±1.21
Sweets, candy	1.42±1.05	1.07±0.83	1.02±0.83	**5.18±2.18**	1.28±1.31
Soft drinks	0.19±0.35	0.17±0.35	0.19±0.37	**1.03±1.18**	0.19±0.47
Fats and oils	1.48±1.15	1.7±1.32	1.11±0.91	**2.06±1.44**	1.74±1.60
Legumes	0.53±0.50	0.48±0.47	0.29±0.47	0.51±0.54	**0.96±0.86**

Figure 2. Profile plot of mean food consumption and cluster grouping for the MrOS cohort.



Selected baseline characteristics of the cohort by cluster membership are shown in **Table 7**. Continuous variables are expressed as mean ± standard deviation, while categorical variables are presented as percentages.

Table 7. Selected baseline characteristics of the MrOS cohort for the 5 cluster solutions

Sample Characteristic	Cluster 1 (n=1006) Bread/pasta	Cluster 2 (n=1364) Light eating	Cluster 3 (n=2848) Many foods	Cluster 4 (n=351) Red meat & fats	Cluster 5 (n=356) Heart healthy	p value*
Age (yr)	73.7±5.8	73.9±6.0	73.7±5.9	73.6±5.7	73.3±5.6	0.36
Height (cm)	174.4±6.9	173.9±6.8	174.1±6.8	173.7±6.6	175.0±6.6	0.07
Weight (kg)	83.3±13.4	82.1±13.0	83.4±13.6	83.3±12.7	84.3±12.1	0.01
Daily Calories (kcal)	1688±658	1584±593	1594±658	1621±608	1717±703	<0.001
Daily Calcium (mg)	1175±579	1141±602	1123±596	1086±553	1188±577	0.025
Daily Vitamin D (IU)	413±242	393±252	384±243	373±248	386±243	0.013
PASE score	149.3±67.3	146.1±68.9	145.4±68.6	158.7±69.6	145.6±65.9	0.53
Drinks/week	4.2±6.9	4.5±7.0	4.2±6.8	3.7±5.8	3.9±6.0	0.18
Body mass index (kg/m^2)	27.4±3.9	27.1±3.7	27.5±3.9	27.6±3.9	27.6±3.5	0.005
Avg walk speed (m/s)	1.22±0.22	1.20±0.23	1.19±0.23	1.19±0.22	1.19±0.22	0.04

Age range n(%)						0.87
64-69	290(28.8)	376(27.6)	838(29.4)	99(28.2)	115(32.3)	
70-74	292(29.0)	396(29.0)	803(28.2)	106(30.2)	99(27.8)	
75-79	245(24.3)	324(23.7)	705(24.7)	87(24.8)	82(23.0)	
80+	179(17.8)	268(19.6)	502(17.6)	59(16.8)	60(16.8)	
Ethnicity n(%)						0.38
White	905(90.0)	1221(89.5)	2532(88.9)	317(90.3)	323(90.7)	
African American	41(4.1)	48(3.5)	132(4.6)	13(3.7)	10(2.8)	
Asian	35(3.5)	50(3.7)	88(3.1)	5(1.4)	10(2.8)	
Hispanic	14(1.4)	32(2.3)	58(2.0)	12(3.4)	9(2.5)	
Other	11(1.1)	13(<1.0)	38(1.3)	4(1.1)	4(1.1)	
Site n(%)						<0.001
Birmingham	212(21.1)	164(12.0)	479(16.8)	68(19.4)	46(12.9)	
Minneapolis	198(19.7)	224(16.4)	471(16.5)	68(19.4)	44(12.4)	
Palo Alto	146(14.5)	270(19.8)	468(16.4)	42(12.0)	70(19.7)	
Pittsburgh	186(18.5)	209(15.3)	441(15.5)	89(25.4)	80(22.5)	
Portland	149(14.8)	255(18.8)	513(18.0)	38(10.8)	52(14.6)	
San Diego	115(11.4)	242(17.7)	476(16.7)	46(13.1)	64(18.0)	
Education n(%)						0.06
<High School	70(7.0)	70(5.1)	190(6.7)	36(10.3)	26(7.3)	
HS, some College	399(40.0)	531(38.9)	1166(40.9)	139(39.6)	158(44.4)	
Completed college	186(18.5)	267(19.6)	500(17.6)	63(18.0)	56(15.7)	
>College	351(35.0)	496(36.4)	992(34.8)	113(32.2)	116(32.6)	
Marital Status n(%)						0.81
Currently Married	828(82.3)	1129(82.8)	2343(83.3)	282(80.3)	288(80.9)	
Other	178(17.7)	235(17.2)	505(16.7)	69(19.7)	68(19.1)	
Smoking status n(%)						0.0045
No	371(36.9)	552(40.5)	1072(37.6)	110(31.3)	116(32.6)	
Past	593(58.9)	777(57.0)	1683(59.1)	222(63.2)	224(62.9)	
Current	42(4.2)	35(2.6)	92(3.2)	19(5.4)	16(4.5)	
Medical History						
Osteoarthritis						0.009
Yes	150(14.9)	254(18.6)	408(14.3)	55(15.7)	54(15.2)	
No	856(85.1)	1110(81.4)	2440(85.7)	296(84.3)	302(84.8)	
Heart Disease**						0.58
Yes	525(52.2)	738(54.1)	1507(52.9)	199(56.7)	194(54.5)	
No	481(47.8)	626(45.9)	1341(47.1)	152(43.3)	162(45.5)	
COPD						0.065
Yes	100(9.9)	134(9.8)	322(11.3)	50(14.2)	31(8.7)	
No	906(90.1)	1230(90.2)	2526(88.7)	301(85.7)	325(91.3)	
Medication Use						
Beta Blocker						0.17
Yes	162(16.1)	248(18.2)	505(17.7)	49(14.0)	71(19.9)	
No	844(83.9)	1116(81.8)	2343(82.3)	302(86.0)	285(80.1)	
ACE inhibitor						0.16
Yes	154(15.3)	244(17.9)	522(18.3)	71(20.2)	68(19.1)	
No	852(84.7)	1120(82.1)	2326(81.7)	280(79.8)	288(80.9)	
Oral hypoglycemic						0.0003
Yes	86(8.5)	112(8.2)	192(6.7)	22(6.3)	47(13.2)	
No	920(91.4)	1252(91.8)	2656(93.3)	329(93.7)	309(86.8)	
Osteoporosis meds						0.08
Yes	21(2.1)	45(3.3)	66(2.3)	6(1.7)	4(1.1)	
No	985(97.9)	1319(96.7)	2782(97.7)	345(98.3)	352(98.9)	
CNS medications***						0.044
Yes	77(7.6)	80(5.9)	201(7.1)	35(10.0)	32(9.0)	
No	929(92.3)	1284(94.1)	2647(92.9)	316(90.0)	324(91.0)	
Statins						0.83
Yes	260(25.8)	340(24.9)	736(25.8)	98(27.9)	89(25.0)	
No	746(74.2)	1024(75.1)	2112(74.2)	253(72.1)	267(75.0)	

Proton pump inhibit.						0.051
Yes	73(7.3)	76(5.6)	202(7.1)	30(8.5)	15(4.2)	
No	933(92.7)	1288(94.4)	2646(92.9)	321(91.4)	341(95.8)	
Loop diuretics						0.011
Yes	40(4.0)	54(4.0)	147(5.2)	16(4.6)	29(8.1)	
No	966(96.0)	1310(96.0)	2701(94.8)	335(95.4)	327(91.8)	
Activities of daily living						
Can walk 2-3 blocks						0.04
Yes	894(88.9)	1218(89.3)	2505(88.0)	293(83.5)	315(88.7)	
No	112(11.1)	146(10.7)	340(11.9)	58(16.5)	40(11.3)	
Can prepare meals						0.49
Yes	986(98.3)	1342(98.5)	2807(98.7)	342(97.4)	350(98.3)	
No	17(1.7)	21(1.5)	38(1.3)	9(2.6)	6(1.7)	

*From GLM, age-adjusted, for continuous characteristics, Pearson Chi-square for categorical characteristics.
**Includes high blood pressure, stroke, angina, and congestive heart failure
***Central nervous system medications Include selective serotonin reuptake inhibitors, benzodiazepines, and non-benzodiazepine anticonvulsants

As shown in Table 6, there were significant differences found between the five cluster groups. For continuous variables, the participants in each cluster differed in their weights ($p=0.01$), amount of daily calories consumed ($p<0.001$), daily Calcium and Vitamin D intake ($p=0.025$ and 0.013, respectively), body mass index ($p=0.005$), and average walking speed ($p=0.04$).

For categorical demographic variables, significant differences were found between clusters and enrollment site ($p<0.0001$). Birmingham had the largest number of men in cluster 1 (21.1%); Palo Alto had the largest number in cluster 2; Portland had the largest number in cluster 3; Pittsburgh had the largest number in both clusters 4 and 5. Smoking status was significantly different across clusters ($p=0.0045$). The most current smokers belonged to cluster 4 (red meat/fats, 5.4%), followed by cluster 5 (heart healthy, 4.5%).

For medical history and medications, significant differences were found across the clusters. There was a significantly greater number of men reporting a history of osteoarthritis (knee, hip, hand) in cluster 2 (18.6%), $p=0.009$. There were no

significant differences across cluster for use of beta-blockers, Ace inhibitors, Cox II inhibitors, corticosteroids (oral or inhaled), statins (HMG CoA reductase inhibitors), or angiotensin II receptor antagonist (not all data shown). However, significant differences across cluster were found for oral hypoglycemic agents (p=0.0003), proton pump inhibitors (p=0.05), loop diruretics (p=0.01), and central nervous system medications (p=0.044).

In examining activities of daily living, there was no significant difference across clusters for the ability to prepare one's own meals (p=0.49). However, the ability to walk 2-3 blocks did significantly differ across clusters (p=0.04). Men in cluster 2 (light eating) had the highest percentage of yes answers (89.3), while men in cluster 4 (red meat and fats) had the lowest percentage of yes answers (83.5).

In age-adjusted analyses, we observed no significant differences in total hip bone mineral density across cluster groupings. **Table 8** presents values for least squares means for bone mineral density across clusters for a model with cluster as the independent variable (adjusted for age). The overall F value for the cluster variable in this model was 1.28 (p=0.27). The R-square value was 0.032, meaning that adjusted for age, the cluster variable explains 3.2% of the variance in total hip BMD. Although no bone densities were statistically significantly different, there was the greatest variation in bone densities in clusters 3 and 5. Further, cluster 5 is the most different from all the other clusters. See **Table 9**.

Table 8. Least square means for total hip bone density in an age-adjusted, simple linear regression analysis		
Cluster	**Total hip BMD LS Mean (g/cm^2)**	**Standard Error**
1.Bread/pasta (n=1006)	0.953	0.0044
2.Light eating (n=1364)	0.953	0.0038
3.Many foods (n=2848)	0.951	0.0026
4.Red meat./fats (n=351)	0.950	0.0074
5.Heart healthy (n=356)	0.969	0.0073

Table 9. p-values for H_0 cluster effect for Table 8 H_0 is LS mean(i)=LS mean(j)					
i/j	**1**	**2**	**3**	**4**	**5**
1		1.0	1.0	1.0	0.37
2	1.0		1.0	1.0	0.29
3	1.0	1.0		1.0	0.17
4	1.0	1.0	1.0		0.39
5	0.37	0.29	0.17	0.39	

The overall F value for cluster for the multivariate model was 2.20, with a marginally significant p-value of 0.07. The R-square value increased considerably to 0.20, meaning that 20% of the variance in total hip bone density was explained by this model. **Table 10** presents the least squares means for total hip BMD when adjusted for all other covariates in the model. **Table 11** presents the differences between clusters, when the null hypothesis is that all LS means are equal to each other. There was a marginally significant difference in total hip BMD between clusters 4 and 5 (p=0.06). No other clusters were significantly different.

Table 10. Least square means for total hip bone density in a multivariate linear regression analysis*		
Cluster	**Total hip BMD LS Mean (g/cm^2)**	**Standard Error**
1.Bread/pasta (n=1006)	0.969	0.006
2.Light eating (n=1364)	0.964	0.006
3.Many foods (n=2848)	0.963	0.005
4.Red meat./fats (n=351)	0.951	0.008
5.Heart healthy (n=356)	0.976	0.008

*other covariates included age, site, race, height, bmi, and self-reported health status.

Table 11. p-values for H_0 cluster effect for Table 10 H_0 is LS mean(i)=LS mean(j)					
i/j	**1**	**2**	**3**	**4**	**5**
1		0.91	0.66	0.15	0.89
2	0.91		0.99	0.40	0.53
3	0.66	0.99		0.47	0.33
4	0.15	0.40	0.47		0.06
5	0.89	0.53	0.33	**0.06**	

After the multivariate model for clusters was completed, we substituted quartiles of factor scores as a categorical variable to examine how prediction of bone mineral density changed as a result of this substitution. **Table 12** lists the overall p-values for each factor score in the multivariate model. Quartiles of Factors 1 (healthy), 4 (diet/low fat), and 6 (salad/tomatoes/fish) were significant (all $p<0.05$).

Table 12. Overall p-values for quartiles of factor score variables in a multivariate linear regression analysis*	
Factor score variable	**p-value**
F1Q	**0.03**
F2Q	0.23
F3Q	0.15
F4Q	**0.047**
F5Q	0.62
F6Q	**0.04**

***other covariates included age, site, race, height, body mass index, and self-reported health status**

Of the three quartiles of factor score variables that were significant in the multivariate model, two of them had significant differences in least squares means between the highest intake (quartile=4) and the lowest intake (quartile=1). For the F1Q variable, those men with the highest factor scores (i.e. those in the 4th quartile) of the healthy factor had a bone density of 0.970, while those in the lowest quartile had a bone density of 0.957 (p for difference=0.02). For the F4Q variable, men with the highest factor scores (i.e. those in the highest quartile) of the diet/lowfat factor had a bone density of 0.972, while those in the lowest quartile had a bone density of 0.959 (p for difference=0.03). See also **Table 13**, below.

Table 13. Least square mean BMD for 2 factors that were significant in the multivariate model.

	Quartiles of Factor 1			**Quartiles of Factor 4**		
	1	4	**p for difference**	1	4	**p for difference**
LS Mean BMD (g/cm^2)	0.957	0.970	0.02	0.959	0.972	0.03

Chapter 4

Discussion

In this cross-sectional analysis of community-dwelling U.S. men ages 65 years and older, we described dietary patterns using two different methods of statistical analysis. In the context of this study's first objective [*use factor and cluster analysis to determine whether different dietary patterns exist among MrOS participants who completed a baseline food frequency questionnaire*], we did derive 6 distinct patterns by factor analysis and 5 distinct groups by cluster analysis. Each factor represents a continuum along which individual men in MrOS vary (via factor scores), while each cluster is a homogeneous subset of individuals based on similarities in food intake. We observed some differences among both factor scores and cluster groupings in relation to selected categorical and continuous variables (**Tables 3, 4, 5, 6, 7**). Men in the highest quartiles of factor 2 (sweets) and factor 3 (western) consumed higher mean daily calories. Men in the highest quartiles of the healthy factor (1) had the highest intakes of calcium and vitamin D. For clusters, we found that subjects had significantly different weights and body mass indices across clusters. Further, the number of men from enrollment site varied by cluster; Birmingham had the highest number of men in Cluster 1 (bread/pasta), while Pittsburgh had the highest numbers of men in both Clusters 4 and 5 (red meat/fats and heart healthy, respectively). It seems likely that dietary patters would tend to vary by site of enrollment, since each area of the country is likely to have regional patterns not found in other sites. This was confirmed by chi-square analysis, where there was a significant difference in cluster groupings by site of enrollment (**Table 7**).

In the context of this study's second objective [*perform a cross sectional analysis to investigate if dietary patterns derived from cluster analysis are associated with bone mineral density of the total hip],* we found a marginally significant relationship between cluster and bone mineral density of the total hip in a multivariate model including age, site, race, body mass index, and self-reported health status. Including other variables, such as medication use and special diet adherence, slightly attenuated the values of the parameter estimates but did not change the direction of the associations. Based on *a priori* decision rules for confounders, it was decided to not include these other variables. Further, we found a marginally significant difference in least-squares means between clusters 4 and 5. Cluster 4 (red meat and fats) had an LS mean total hip BMD of 0.951, compared to 0.976 for Cluster 5 (heart healthy). Although this difference may seem small in magnitude, it accounts for nearly one quarter of the standard deviation in measurement of total hip BMD. This difference could be clinically significant, especially if small changes in micronutrient intakes between clusters accounts for the differing BMD values. Further, our findings are consistent with another study that used cluster analysis to examine bone density. Tucker et al found that men in a cluster with high servings of fruits, vegetables, and cereals (similar to our Cluster 5) had a significantly higher BMD measured at Ward's area and femoral neck than did men in a "candy" cluster (2002). It is interesting to note that our cluster 4, red meat and fats, also had the highest mean servings of candy, much like the cluster with the lowest BMD in Tucker's 2002 study.

Few studies have used bone mineral density as the outcome in analysis of dietary patterns. Tucker et al observed 6 distinct dietary patterns using the Framingham cohort of elderly adults. Their cohort had similar baseline BMI measurements (27.1)

but a higher percentage of current smokers (9.3%) than did men participating in MrOS. Men in the fruit, vegetable, and cereal group had significantly higher bone mineral density than did men in the other groups (2002). We found that the least squares mean difference between the red meat and fats cluster versus the heart healthy cluster was marginally significant (p=0.06). Differences in results may be explained by differences in individual samples, analytic choices by the investigator, or differences in food frequency questionnaires. As noted by Jacques et al, it is precisely a lack of population specificity in describing dietary patterns that can make it difficult to compare studies using that approach (2001).

Of all the clusters, cluster 4 (red meat and fats) had the lowest total hip BMD measurement. There are plausible biological reasons this could occur. Protein intake has been previously associated with urinary loss of calcium (Heaney, 2001). Studies have not been as clear when trying to establish a link between high protein intake and loss of bone mineral density. Hannan et al observed that total protein and animal protein intakes were protective for BMD levels in the Framingham cohort of elderly adults (2000). Conversely, Sellmeyer et al observed that the ratio of animal to vegetable protein was important for bone density; those with higher ratios had greater bone loss at the femoral neck. However, those individuals with relatively lower intake of protein had increased bone loss, even after controlling for confounders (2001). In addition, only women were examined in this study; it is possible that men might have different results and may not be comparable to women. As suggested by Heaney, it is difficult to resoundingly prove "skeletal harm" related to intake of protein, either animal or vegetable (2001). Cluster 4 also had the lowest intakes of both calcium and Vitamin D, which could alternately explain the low

bone density values for men in that cluster. When we controlled for intakes of both of these substances, the difference in LS means between clusters 4 and 5 was no longer marginally significantly different. However, the cluster variable represents dietary intake derived from the food frequency questionnaire; we must be cautious that this result is not affected by statistical over adjustment.

Substitution of Quartiles of Factor Scores into the Multivariate Model

When we substituted quartiles of factor scores into the multivariate model constructed for the cluster analysis, we found that 3 factors (1, healthy; 4, diet/lowfat; and 6, Mediterranean) were significantly associated with bone density of the total hip (all $p < 0.05$). When LS means were examined, it was found that the differences between the lowest and highest quartiles of factors 1 and 4 were significantly different. It should be noted that differences in the results when using quartiles of factor scores instead of clusters are probably expected, since these are two methodologically different procedures. Factor analysis creates patterns of food intake based on the correlations between foods; every person receives a score for each pattern extracted. These factors are *not* mutually exclusive. Conversely, cluster analysis creates patterns that *are* mutually exclusive that are defined by maximizing the differences in food intakes (food groups). It might be easier to interpret results of cluster analysis, since each person belongs to only one cluster, and each cluster has a specific composition of foods and nutrients (Newby et al, 2004).

Study Limitations

This investigation does have limitations. The current analysis used food frequency questionnaire data to assess diet. This retrospective method of diet assessment is subject to measurement error. This may lead to misclassification of respondents into dietary exposure categories. Thus, if the FFQ inaccurately characterizes dietary differences between diseased and non-diseased persons, there will be a reduction in the ability to identify dietary risk factors. This does not seem likely to affect the validity of the current study, since men did not know their BMD values at time of enrollment. Even though the Block FFQ has been validated, there are inherent limitations to this form of dietary analysis; some limitations include exclusion of certain foods on the FFQ, and seasonal variations in diet that may affect the precision of the dietary measurements. Finally, the generalizability of the MrOS study may be limited. The cohort of men is highly educated and a small number of the men are current smokers. Thus these men may be dissimilar to the U.S. population of men over the age of 65 years.

Some of these limitations may explain why little differences were observed among key characteristics across factor scores. Although no other studies were found that used factor analysis to describe bone mineral density, many other chronic disease outcomes have been used. In the current study, we found that body mass index values varied little across quartiles of any of the factor scores. Participants in the highest quartile of factor 1 (healthy) had a BMI of 27.2±3.7, while subjects in the lowest quartile of factor 1 had a BMI of 27.4±3.8. Newby et al found a more striking difference in BMI level across quintiles of factor scores for their factor 1 (reduced fat

dairy products, fruit, and fiber); subjects in the highest quintile of factor 1 had a lower BMI than those in the lowest quintile of factor 1. Instead of bone mineral density, the outcome in that study was anthropometric changes (2004). A study by Hu et al using factor analysis to describe risk of coronary heart disease in men found that men in the highest quintile of the "prudent" dietary pattern had a BMI of 25±3, while men in the lowest quintile had a BMI of 26±3 (2000). Possible reasons for lack of variation in the current study could be measurement of food intake (Newby used a 7-day dietary record; Hu used a 131-item food frequency questionnaire). Although the studies described above tend to have at least one similar pattern (heart healthy, prudent, low-fat dairy and grains), it remains difficult to reproduce results from any one study (Jacques et al, 2001).

Factor analysis has been criticized for a number of reasons. An important decision in factor analysis is the choice of the number of factors to be extracted from the data. There are various methods for these decisions, but all involve subjective decision making on the part of the investigator. Usually, a factor is considered to be important if its eigenvalue is greater than 1.0 (i.e. if the factor explains more of the variance in the correlations than is explained by a single variable). However, other authors have used eigenvalues of >1.25 if the number of factors was too great for further analysis (Slattery et al, 1998; Martinez et al, 1998). The method of factor rotation (which redefines factors in order to make sharper distinctions in the meaning of factors) also varies. It is important to remember that methods of rotation do not necessarily improve the fit of the data; instead they redistribute the explained covariance among the factors (Martinez et al, 1998). The naming of the factors can lead to bias, especially if the variance explained by the name is fairly low. Users of

this technique should be aware that naming of factors is more art than science (Kachigan, 1986). However, publishing the actual factor loadings can help to allow each reader to make her own interpretations (Slattery et al, 1998). Finally, a first step in performing factor analysis with dietary data is often to create food groupings from the food items on the FFQ. A review of current literature has revealed that most methods for this step involve using similarity of macronutrient content or groupings from previous studies (Tucker et al 2002; Appel et al, 1997; Kant, 2004; Quatromoni et al, 2002; Slattery et al, 1998; Millen et al, 1996). The latter method is done to limit subjectivity in defining groups and for comparisons with other studies (Kerver et al, 2003). However, the approach in this analysis is unique in that no *a priori* groupings were made. This approach is strictly data-driven, and allows for groupings that may not be usual or expected.

Cluster analysis also involves some subjective decisions. Unlike factor analysis, individual foods from the FFQ must first be placed into a smaller number of groups (using previous studies or existing knowledge). Cluster analysis is sensitive to outliers, so each investigator should have a way to exclude implausible daily food consumption. Also, the number of clusters is not predefined. The investigator must first specify a number to try in the analysis, and then select a final cluster set by comparing between and within cluster variance, computing Scree plots, and examining the clusters for nutritionally meaningful separation (Tucker et al, 2002).

Even with the limitations described above for factor and cluster analyses, these approaches are likely to work well within a specific population for identifying dietary patterns found in that sample. Although the results may not be necessarily

applicable to other populations, the internal validity of the current study should be maintained. However, the results of this study may not be comparable to other studies; the external validity may be affected due to subjective decisions required by this type of analysis.

Finally, the cross-sectional nature of the analysis of dietary patterns and BMD cannot discern a cause and effect relationship. However, this information will still be important in determining public health guidelines, assessing the adequacy of the current dietary recommendations, and hypothesis generation.

Strengths

This analysis has strengths as well. It provides an alternative approach for studying food patterns in older men. By not developing *a priori* groupings of foods before starting the statistical procedures, unusual or unexpected patterns may be discerned. The dietary pattern approach may additionally reflect the multidimensionality of dietary behaviors in this cohort. Although population specificity makes it hard to compare patterns across different studies (i.e. the patterns may not reflect ideal diets), the patterns found do in fact reflect the patterns in the population under study (Jacques et al, 2001). Therefore, they provide useful information for health promotion.

Chapter 5

Summary and Conclusions

In this cross sectional analysis of dietary patterns among men ages 65 years and older, we found that different dietary patterns could be determined using factor and cluster analysis; we found six distinct patterns using factor analysis, and five distinct patterns using cluster analysis. Diet is only one important factor in determining bone health. We found that men belonging to the cluster of red meat and fats had the lowest bone mineral density measurement at the total hip. We did not find an overall significant association between dietary pattern and bone mineral density. However, results suggest that men with higher intakes of "healthy" foods (e.g. more servings of fruits, vegetables, and grains) have higher bone densities. Although cause and effect cannot be ascertained in this type of analysis, the results shown here may have implications for dietary recommendations for men in this age group.

Future studies should use a dietary patterns approach to study factors important for fracture etiology, rather than only focusing on bone mineral density. Researchers should continue to improve and refine the methodology for both factor and cluster analysis when using these tools to examine dietary data. Examining the holistic pattern of eating should complement the research conducted using single foods or nutrients. These methods should cement the knowledge that the recommended dietary guidelines have positive health effects, and perhaps provide hypotheses for further research.

REFERENCES

Appel LJ, Moore TJ et al. A Clinical Trial of the Effects of Dietary Patterns on Blood Pressure. NEJM 1997; 336(16):1117-24.

Assessment of fracture risk and its application to screening for post-menopausal osteoporosis: report of a World Health Organization Study Group. WHO Technical Report Series No. 843. Geneva: WHO 1994.

Berkeley Nutrition Services References (no date). Available at http://www.lib.umich.edu/handouts/apa.pdf. Retrieved on August 17, 2004.

Blank JB, Cawthon PM et al. Overview of recruitment for the osteoporotic fractures in men study (MrOS). Contemp Clin Trials 2005;26(5):557-68.

Booth SL, Tucker KL et al. Dietary vitamin K intakes are associated with hip fracture but not with bone mineral density in elderly men and women. Am J Clin Nutr 2000; 71:1201-08.

Cauley JA, Fullman FL et al. Factors Associated with Lumbar Spine and Proximal Femur Bone Mineral Density in Older Men. Osteoporosis International (epub online; 11 May 2005).

Cauley JA, Zmuda JM 1999. Risk Factors For Fractures in Men. In: Orwoll E (ed.) Osteoporosis in Men: The Effects of Gender on Skeletal Health. Academic Press, San Diego, CA, pp. 363-393.

Dawson-Hughes B, Dallal GE et al. A controlled trial on the effect of calcium supplementation on bone density in postmenopausal women. N Engl J Med 1990; 323:878-83.

Dawson-Hughes B, Dallall GE et al. Effect of Vitamin D Supplementation on Wintertime and Overall Bone Loss in Healthy Postmenopausal Women. Ann Int Med 1991; 115(7): 505-512.

de Lorgeril M, Salen P et al. Mediterranean Dietary Pattern in a Randomized Trial: Prolonged Survival and Possible Reduced Cancer Rate. Arch Int Med 1998; 158(11):1181-1187.

Fox Ray N, Chan JK et al. Medical Expenditures for the Treatment of Osteoporotic Fractures in the United States in 1995: Report from the National Osteoporosis Foundation. J Bone Min Res 1997; 12(1):24-35.

Garson, GD. Cluster Analysis. Available at http://www2.chass.ncsu.edu/garson/pa765/cluster.htm. Retrieved October 10, 2004.

Hannan MT, Tucker KL et al. Effect of Dietary Protein on Bone Loss in Elderly Men and Women: The Framingham Osteoporosis Study. J Bone Miner Res 2000;15(12):2504-2512.
Heaney RP. Protein intake and bone health: the influence of belief systems of the conduct of nutritional science. Am J Clin Nutr 2001;73:5-6.

Heaney RP. The role of nutrition in prevention and management of osteoporosis.Clin Obs Gyn 1987; 50:833-846.

Hu F. Dietary pattern analysis: A new direction in nutritional epidemiology. Curr Opin Lipidol 2002; 13:3-9.

Hu FB, Stampfer MJ et al. Prospective study of major dietary patterns and risk of coronary heart disease in men. Am J Clin Nutr 2000; 72:912-21.

Jacques PF, Tucker KL. Are dietary patterns useful for understanding the role of diet in chronic disease? Am J Clin Nutr 2001;73:1-2.

Johnell O, Kanis J. Epidemiology of osteoporotic fractures. Osteoporosis Int 2004; September 8 (Epub ahead of print). Available at http://www.springerlink.com/app/home/contribution.asp?wasp=fcgnngmglg31e6tvxceg&referrer=parent&backto=issue,35,102;journal,1,105;linkingpublicationresults,1:102828,1).

Kachigan, SK. Statistical Analysis: An Interdisciplinary Introduction to Univariate and Multivariate Methods. Radius Press; New York, 1986.

Kant AK, Graubard BI et al. Dietary Patterns Predict Mortality in a National Cohort: The National Health Interview Surveys, 1987 and 1992. J Nutr 2004; 134:1793-99.

Kant AK, Schatzkin A et al. A Prospective Study of Diet Quality and Mortality in Women. JAMA 2000; 283(16):2109-15.
Kant AK. Dietary Patterns and Health Outcomes. J Am Diet Assoc 2004; 104:615-635.

Kerver JM, Yang EJ et al. Dietary patterns associated with risk factors for cardiovascular disease in healthy US adults. Am J Clin Nutr 2003; 78:1103-10.

Kim MK, Sasaki S et al. Prospective study of three major dietary patterns and risk of gastric cancer in Japan. Int J Cancer 2004; 110:435-42.

Koffmann K, Schulze MB et al. Dietary patterns: report of an international workshop. Pub Health Nutr 2000; 5(1):89-90.

Lewis, Carol. Osteoporosis and Men. FDA Consumer Magazine 2002; September-October. U.S. Department of Health and Human Services, U.S. Food and Drug Administration. Available online at

(http://www.fda.gov/fdac/features/2002/502_men.html. Retrieved October 3, 2004.

Martinez ME, Marshall JR et al. Invited Commentary: Factor Analysis and the Search for Objectivity. Am J Epi 1998; 148(1):17-19.

McCann SE, Marshall JR et al. Analysis of patterns of food intake in nutritional epidemiology: food classification in principal components and the subsequent impact on estimates for endometrial cancer. Pub Health Nutr 2001; 4(5):989-997.

McCann SE, Weiner J et al. Is principal components analysis necessary to characterize dietary behavior in studies of diet and disease? Pub Health Nutr 2001l 4(4):903-908.

Millen BE, Quatromoni PA et al. Dietary Patterns of Men and Women Suggest Targets for Health Promotion: The Framingham Nutrition Studies. Am J Health Prom 1996; 11(1):42-53.

National Health and Nutrition Examination Survey, Data Tables, Anthropometric Reference Data, United States, 1988-1994, Table 5. Available at http://www.cdc.gov/nchs/about/major/nhanes.

National Osteoporosis Foundation. Fast Facts (no date). Retrieved on October 4, 2004. Available at http://www.nof.org/osteoporosis/diseasefacts.htm.

Newby PK, Hallfrisch J et al. Food patterns measured by factor analysis and anthropometric changes in adults. Am J Clin Nutr 2004; 80:504-13.

Newby PK, Muller D et al. Associations of empirically derived eating patterns with plasma lipid biomarkers: a comparison of factor and cluster analysis methods. Am J Clin Nutr 2004; 80(3):759-67.

Newby PK, Tucker KL. Empirically Derived Eating Patterns Using Factor or Cluster Analysis: A Review. Nutr Reviews 2004; 62(5):177-203.

Obrant KJ, Bengner U et al 1989. Increasing age-adjusted risk of fragility fractures: A sign of increasing osteoporosis in successive generations? In: Orwoll E (ed.) Osteoporosis in Men: The Effects of Gender on Skeletal Health. Academic Press, San Diego, CA, pg. 363.

Orwoll E, Blank JB et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study—a large observational study of the determinants of fracture in older men. Contemp Clin Trials 2005; 26(5):569-85.

Osteoporosis Prevention, Diagnosis and Treatment. NIH Consensus Statement 2000; 17(1): March 27-29. U.S. Department of Health and Human Services, National Institutes of Health, Bethesda, MD. Available at

http://consensus.nih.gov/cons/111/111_statement.pdf. Retrieved October 17, 2004.

Pohlmann JT. Factor Analysis Glossary (no date). Available at http://www.siu.edu/~epse1/pohlmann/factglos/.

Quatromoni PA, Copenhafer DL et al. The internal validity of a dietary pattern anaysls: The Framingham Nutrition Studies. J Epi Comm Health 2002; 56:381-388.

SAS Institute. SAS User's guide, version 9.1 Cary, NC: SAS Institute, 2002-2003.

Schulze MB, Hoffmann K et al. Risk of Hypertension among Women in the EPIC-Potsdam Study: Comparison of Relative Risk Estimates for Exploratory and Hypothesis-oriented Dietary Patterns. Am J Epi 2003; 158(4):365-73.
Sellmeyer DE, Stone KL et al. A high ratio of dietary animal to vegetable protein increases the rate of bone loss and the risk of fracture in postmenopausal women. Am J Clin Nutr 2001;73:118-22.

Slattery MK, Boucher KM et al. The Senior Authors' Response: Factor Analysis as a Tool for Evaluating Eating Patterns. Am J Epi 1998; 148(1): 20-21.

Slattery ML, Boucher KM et al. Eating Patterns and Risk of Colon Cancer. Am J Epi 1998; 148(1):4-16.

Tucker KL, Chen H et al. Bone mineral density and dietary patterns in older adults: the Framingham Osteoporosis Study. Am J Clin Nutr 2002; 76:245-52.

Tucker KL, Hannan MT et al. Potassium, magnesium, and fruit and vegetable intakes are associated with greater bone mineral density in elderly men and women. Am J Clin Nutr 1999; 69:727-36.

University of Texas at Austin Statistical Services. Factor Analysis Using SAS PROC FACTOR. 1995-1997. Available at http://www.ats.ucla.edu/stat/sas/library/factor_ut.htm.

U.S. Department of Health and Human Services. Administration on Aging. Statistics on the Aging Population. Available at http://www.aoa.gov/prof/Statistics/statistics.asp. Revised September 2004. Retrieved 2 October 2004.

U.S. National Research Council. Committee on Diet and Health: Implications for Reducing Chronic Disease Risk. National Academy Press; Washington DC 1989.

Wirfalt AK, Jeffery RW. Using cluster analysis to examine dietary patterns: Nutrient intakes, gender, and weight status differ across food pattern clusters. J Am Diet Assoc 1997; 97(3):272-79.

Wirfalt E, Mattisson I et al. Food patterns defined by cluster analysis and their utility as dietary exposure variables: a report from the Malmo Diet and Cancer Study. Pub Health Nutr 1999; 3(2):159-173.

Appendix A
Foods used in the dietary pattern analyses

- **Breakfast-style foods**
  - Eggs (including egg biscuits, Egg McMufffins) NOT egg substitute
  - Bacon or breakfast sausage, including sausage biscuit
  - Pancakes, waffles, or French toast
  - Cooked cereals like oatmeal, cream of wheat or grits
  - Cold cereals like corn flakes, cheerios, Special K, fiber cereals
  - Cheese, sliced or spread, including on sandwiches
  - Low fat cheese
  - Yogurt or Frozen Yogurt
- **Fruits**
  - Bananas
  - Fresh apples or pears
  - Oranges, tangerines, not including juice
  - Applesauce, fruit cocktail, or any canned fruit
  - Any other fruit like grapes, honeydew, pineapple, strawberries
- **Vegetables, fresh, frozen, canned, or in stir fry, at home or in restaurant**
  - French fries, fried potatoes, hash browns
  - White potatoes not fried, including boiled, baked, mashed, potato salad
  - Sweet potatoes, yams
  - Rice, or dishes made with rice
  - Baked beans, chili with beans, blackeye peas, any other dried beans
  - Corn
  - Green beans or green peas
  - Broccoli
  - Carrots, or stews or mixed vegetables containing carrots
  - Spinach, or greens like collard
  - Cole slaw, cabbage
  - Green salad
  - Raw tomatoes, including in salad
  - Salad dressing
  - Low-fat salad dressing
  - Any other vegetables, like okra, cooked green peppers, cooked onions
  - Tofu, bean curd
  - Vegetable soup, vegetable beef, chicken vegetable, or tomato soup
  - Other soups, like chicken noodle, chowder, mushroom, instant soups
- **Meats**
  - Hamburgers, cheeseburgers, meat loaf, at home or in a restaurant
  - Beef steaks, roasts, pot roast, or in frozen dinners or sandwiches
  - Liver, including chicken livers or liverwurst
  - Pork, including chops, roasts, or dinner ham
  - Beef or pork trimmed of fat
  - Mixed dishes with meat or chicken, like stew, corned beef hash
  - Fried chicken, at home or in a restaurant
  - Chicken or turkey not fried, such as baked, grilled, or on sandwiches
  - Chicken without skin
  - Shellfish like shrimp, scallops, or crab
  - Fish or fish sandwich at home or in a restaurant

Hot dogs, or sausage like Polish, Italian, or Chorizo
Boloney, sliced ham, turkey lunch meat, other lunch meat
Low-fat lunch meats

- **Pasta, breads, spreads, snacks**

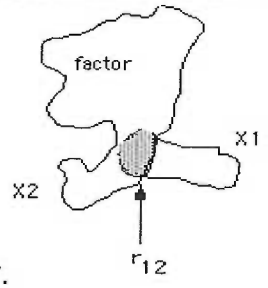
Spaghetti, lasagna, or other pasta with tomato sauce
Cheese dished without tomato sauce
Pizza, including carry out
Biscuits, muffins
Rolls, hamburger buns, English muffins, bagels
White bread or toast, including French, Italian or in sandwiches
Dark bread like rye or whole wheat, including in sandwiches
Margarine in cooking, or on bread, potatoes, or vegetables
Butter in cooking, or on bread, potatoes, or vegetables
Mayonnaise, sandwich spreads
Peanut butter
Gravy
Snacks like potato chips, corn chips, popcorn (not pretzels)
Peanuts, other nuts or seeds
Crackers
Doughnuts, cake, pastry
Cookies
Low-fat cookies
Ice cream, ice milk, ice cream bars
Low-fat ice cream
Pie or cobbler
Chocolate candy, candy bars

- **Beverages**

Real 100% orange or grapefruit juice, including fresh, frozen, bottled
Calcium-fortified orange juice
Hi-C, Kool-aid, or other drinks with added Vitamin C
Tomato juice or V-8 juice
Liquid supplements like Ensure, Instant breakfast, or diet shakes
Glasses of milk, any kind
Milk choices: Whole, 2%, 1%, skim, rice, soy
Soft drinks with caffeine, like colas or Mountain Dew
Coffee
Decaffeinated coffee
Tea, regular black or Chinese (not herbal)
Additives to tea and coffee: Cream/half & half, nondairy creamer, milk

Appendix B: Selected statistical definitions (Pohlmann J, no date)

Common factor analysis: a statistical technique that uses the correlations between observed variables to estimate common factors and the structural relationships linking factors to observed variables. The diagram below illustrates how two observed variables can correlate because of their relationships with a common factor.



Cluster analysis: a collection of statistical techniques for creating homogeneous groups of cases or variables. Clusters are formed using distance functions. The elements in a cluster have relatively small distances from each other and relatively larger distances from elements outside of a cluster. See *distance.*

Distance: a measure of the disparity between two observations on a set of variables. The most common measure is the squared Euclidian distance, which is the sum of squared differences across a set of variables. Letting m = the number of variables, and Xij be the value of the j-th variable for the i-th case, the squared Euclidian distance between cases k and l is

$$D_{kl}^2 = \sum_{j=1}^{m} (Xkj - Xlj)^2$$

Distance functions are used in cluster analysis to form clusters of variables or cases that are most similar or have small distances.

Eigenvalue: the variance in a set of variables explained by a factor or component, and denoted by **lambda**. An eigenvalue is the sum of squared values in the column of a factor matrix, or

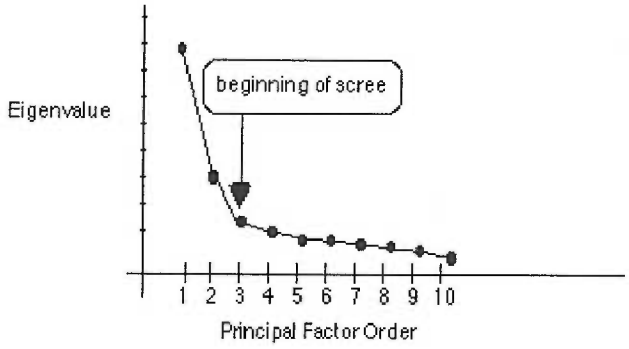
$$\lambda_k = \sum_{i=1}^{m} a_{ik}^2$$

where a_{ik} is the factor loading for variable i on factor k, and m is the number of variables.

Orthogonal decomposition of variables: transforming a set of correlated variables into a set of uncorrelated variables.

Principal components analysis: (1) a method of factoring a correlation matrix directly, without estimating communalities. Linear combinations of variables are estimated which explain the maximum amount of variance in the variables. The first component accounts for the most variance in the variables. Then the second component accounts for the most variance in the variables residualized for the first component, and so on. (2) transforms a collection of measured variables into a set of orthogonal maximum variance linear combinations.

Scree test: a graphic method for determining the number of factors. The eigenvalues are plotted in the sequence of the principal factors. The number of factors is chosen where the plot levels off to a linear decreasing pattern. The figure below suggests a two-factor solution, since the eigenvalues begin a linear decline commencing with the third factor.



Varimax rotation: an orthogonal rotation criterion that maximizes the variance of the squared elements in the columns of a factor matrix. Varimax is the most common rotational criterion.