Self-Reported Medical Conditions and Demographic, Behavioral and Dietary Factors Associated with Serum 25(OH)-Vitamin D Concentration in the US Adult Population

Self-Reported Medical Conditions and Demographic, Behavioral and Dietary Factors Associated with Serum 25(OH)-Vitamin D Concentration in the US Adult Population

by

WILLIAM E. VAN FLEIT

Under the Direction of Xu Zhang

ABSTRACT

This research uses data from the 2003-2006 National Health and Nutrition Examination Survey (NHANES) to determine dietary and other factors associated with serum 25(OH)-Vitamin D concentration for 5,474 adults age 20 years and older. After multivariate adjustment, we found that serum 25(OH)-Vitamin D concentration was positively associated with diets high in fruits, vegetables, and lean meats, while diets high in processed foods and high-fat meats were inversely associated with vitamin D level. Serum 25(OH)-Vitamin D concentration was also significantly associated with age, gender, race/ethnicity, BMI, physical activity, supplementation, and the season of survey administration. Self-reported cardiovascular and kidney disease were significantly associated with serum 25 (OH)-Vitamin D concentration after adjustment for significant confounders.

INDEX WORDS: Centers for Disease Control and Prevention, CDC, Chronic Disease, Dietary Profile, National Health and Nutrition Examination Survey, NHANES, Nutrition, Principal Components Analysis, Vitamin D

Self-Reported Medical Conditions and Demographic, Behavioral and Dietary Factors Associated with Serum 25(OH)-Vitamin D Concentration in the US Adult Population

by

WILLIAM E. VAN FLEIT

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2012

Self-Reported Medical Conditions and Demographic, Behavioral and Dietary Factors Associated with Serum 25(OH)-Vitamin D Concentration in the US Adult Population

by

WILLIAM E. VAN FLEIT

Committee Chair: Xu Zhang

Committee: Vijay Ganji

Yuanhui Xiao

Yichuan Zhao

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2012

ACKNOWLEDGEMENTS

Many thanks to Dr. Xu Zhang, Dr. Vijay Ganji, Dr. Yuanhui Xiao, Dr. Yichuan Zhao and Bernadette Martineau. I would also like to offer special thanks to my wife, Callie, for her support throughout this project.

TABLE OF CONTENTS

LIST OF TABLES

1 INTRODUCTION

1.1 Background

Knowledge of the role of vitamin D in the human body and its relationship to disease has been of key public health importance since the discovery of the relationship between vitamin D deficiency and rickets in the early twentieth century (Rajakumar et al., 2007). Vitamin D is now understood to play an important role in many bodily processes, including calcium absorption, neuromuscular function and inflammation response (Norman, 2008). Recently published research has shown an association between vitamin D deficiency and a number of chronic diseases such as diabetes, cancer and cardiovascular disease (Mitri 2011; Ma et al., 2011; Leu et al., 2010).

Our research uses data from the National Health and Nutrition Examination Survey (NHANES) to study the determinants of serum vitamin D concentration, and the relationship between vitamin D and self-reported medical conditions. We consider several key demographic, behavioral and environmental variables in our analysis. We also use principal components analysis to determine three major dietary patterns based on dietary recall data from the NHANES Food Frequency Questionnaire. These dietary patterns are used as the primary variables for determining serum concentration of vitamin D, and are analyzed together with the other variables to determine the multivariate-adjusted association between vitamin D insufficiency or deficiency and five self-reported medical conditions.

1.2 Sources of Vitamin D

Most of the vitamin D that is used by the human body is produced in the skin as a result of exposure to ultraviolet radiation. Few unfortified foods contain nutritionally significant levels of vitamin D. The flesh and oils of fatty cold water fish, and to a lesser extent egg yolks, cheese and beef liver, are among those foods that do contain significant levels (U.S. Department of Agriculture, 2012). However, most vitamin D enters the diet through fortification of foods such as milk, orange juice and breakfast cereals (Rajakumar et al., 2007). Vitamin D may also be obtained through supplementation.

There are two chemically different forms of vitamin D that are used by the human body (Gallieni et al., 2009). The first form, vitamin D_3, is called cholecalciferol. Vitamin D_3 makes up approximately 90% of the serum concentration of vitamin D circulating in the bloodstream (Lehman and Meurer, 2010). Although vitamin D_3 may be obtained from dietary sources, under usual circumstances most of the body's supply is produced endogenously in the skin. The production of vitamin D_3 in the skin is a multistep process that begins when 7-dehydrocholesterol, or provitamin D_3, is exposed to ultraviolet light and transformed into previtamin D_3. Vitamin D_3, cholecalciferol, is formed through a heat dependent isomerization of previtamin D_3 that takes place over several hours after the formation of the previtamin D_3 (Gallieni et al., 2009).

The second from of vitamin D is ergocalciferol, or vitamin D_2. This is a synthetic form of vitamin D that is produced when ergosterol, found naturally in fungal cell membranes, is irradiated with ultraviolet light (Gallieni et al., 2009). Vitamin D_2 is the source of vitamin D that is used to fortify foods such as milk, orange juice and breakfast cereal. Vitamin D_2 is also the source of vitamin D in many dietary supplements (Gallieni et al., 2009). Though a major source of dietary vitamin D, researchers have shown that vitamin D_2 is only one-third as effective in raising total serum vitamin D concentrations as compared to vitamin D_3 (Norman, 2008).

Vitamin D_2 and vitamin D_3 are converted by the liver into 25-hydroxyvitamin D_2 and 25-hydroxyvitamin D_3, respectively, which then circulates in the bloodstream. These two prohormones may be referred to collectively as 25-hydroxyvitamin D, or 25(OH)D. Circulating 25(OH)D is then converted primarily by the kidneys, but also by other types of cells, into the hormone 1,α,25 dihydroxyvitamin D, or 1,α,25(OH)D (Lehman and Meurer, 2010). This hormone has receptors throughout the body, the role of which in human health and disease is the topic of much recent scientific research (Makariou et al., 2011).

1.3 Measuring Vitamin D Status

Nutritional studies require a measurement of the level of vitamin D in the body that reflects the adequacy of dietary intakes and environmental exposures needed to produce vitamin D. While the steroid hormone 1,α,25(OH)D is known to be the physiologically active metabolite of vitamin D, serum concentrations of this hormone are tightly regulated by the body's endocrine system (Norman, 2008). Thus, the serum concentration of 1,α,25(OH)D is not necessarily a good measure for investigating the adequacy of vitamin D-related nutritional intakes and environmental ultraviolet light exposure. For this reason, measurements of the circulating concentration of the prohormone 25(OH)D are commonly used to determine whether a person has adequate nutritional or behavioral prerequisites for a healthy level of vitamin D.

The serum 25(OH)D concentration is measured through the use of a laboratory test, known as a 25(OH)D assay, that is conducted using a serum specimen. One version of the test involves chemically extracting 25(OH)D from serum, and then measuring the serum concentration through the use of an antibody that binds specifically to 25(OH)D (CDC, 2003). Several manufacturers produce the 25(OH)D assay, and actual serum concentration values obtained by different versions of the assay may not be directly comparable due to underlying differences between assays. 25(OH)D concentration is normally expressed in either nanomoles per liter (nmol/L) or nanograms per milliliter (ng/mL). These two measures of concentration may be converted using the relationship 1 ng/mL = 2.496 nmol/L (Ginde et al., 2009).

1.4 Vitamin D Deficiency

There have been attempts by various professional or governing bodies to define an adequate level of vitamin D based on specific criteria or risk of disease. Despite these attempts, there are no universally accepted guidelines regarding the levels of vitamin D that are optimal for good health. The Institute of Medicine (2010) has published guidelines to establish vitamin D adequacy based on serum 25(OH)D concentration. According to these guidelines, vitamin D deficiency is defined by serum 25(OH)D concentration less than 12 ng/mL. This minimum threshold is based on the levels associated with increased risk of rickets in children and osteomalacia in adults. Serum 25(OH)D concentrations between 12 and 20 ng/mL are considered to be "inadequate" based on evidence associated with bone health and overall health. Serum 25(OH)D concentrations above 20 ng/mL are considered to be "adequate." The Institute of Medicine (2010) has noted evidence that high 25(OH)D intakes (above 4,000 IU per day) may be associated with an increased risk of cardiovascular disease, and that very high intakes (above 10,000 IU per day) have been associated with tissue damage.

The prevalence of vitamin D deficiency and associated risk factors have been the topic of much research. Overall, vitamin D deficiency appears to be widespread in the United States and throughout the rest of the world. Using data from the National Health and Nutrition Examination Survey for the years 2001 through 2006, Ganji, Zhang and Tangpricha (2012) found that the prevalence of vitamin D inadequacy or deficiency (serum 25(OH)D concentration $\leq$ 20 ng/mL) affected approximately 32% of the U.S. adult population. Generally, the risk of vitamin D deficiency, regardless of cutpoints, tends to increase with age and is highest among the elderly (Forrest and Stuhldreher, 2011; Ginde et al., 2009). Mansbach, Ginde and Carmago (2009) found that the prevalence of similarly defined vitamin D inadequacy or deficiency among children aged 1 to 11 years old was 18%. Regardless of age, researchers have consistently identified persons of female gender and those with darker complexions (non-White race/ethnicity) as having increased risk of vitamin D deficiency (Forrest and Stuhldreher, 2011; Ginde et al., 2009, Kumar et al., 2009; Mansbach, Ginde and Carmago, 2009) .

1.5 Vitamin D and Health

1.5.1 Cancer

The link between cancer incidence and vitamin D deficiency is not clear. The proposed anti-cancer activity of vitamin D stems from its role in cellular proliferation and metabolism, as well as its involvement in inhibiting inflammation (Mocellin, 2011). While some observational studies suggest that a relationship between cancer prevalence and vitamin D deficiency exists, Mocellin (2011) points out that clinical trials provide less conclusive evidence. It is possible that the anti-cancer activity of vitamin D may be most relevant for certain types of cancers. In a meta-analysis across nine observational studies, Ma et al. (2011) found that higher vitamin D levels were associated with lower risk of developing colorectal cancer (RR 0.67, 95% CI 0.54 to 0.80). Freedman et al. (2010) did not find a consistent association between serum vitamin D concentration and cancer mortality.

1.5.2 Cardiovascular Disease

There is evidence that vitamin D may be an important factor in the prevention of cardiovascular disease. Zittermann, Schleithoff and Koerfer (2005) suggest several ways that vitamin D may affect the cardiovascular system by reducing vascular calcification. Vitamin D may act through vitamin D receptors to suppress the formation of vascular smooth muscle cells, which in turn reduces the uptake of calcium into these cells. Vitamin D is strongly positively associated with the production of matrix Gla protein, which is associated with reduced calcification. Vitamin D may also reduce the risk of cardiovascular disease through its anti-inflammatory action and by inhibiting parathyroid hormone. Research suggests that the risk of cardiovascular disease is most pronounced at serum vitamin D levels below 15 ng/mL (Leu and Giovannucci, 2011). Dietary intakes of vitamin D of greater than 600 IU/day have also been found to be associated with lower risk of cardiovascular disease (RR 0.84, 95% CI 0.72 to 0.97) (Sun et al., 2011).

1.5.3 Diabetes

There is evidence of an association between vitamin D deficiency and the risk of diabetes. In a meta-analysis of four cohort studies, Mitri, Muraru and Pittas (2011) found that higher serum vitamin D concentrations were associated with a lower risk of type 2 diabetes. Persons with a serum a 25(OH)D concentration of at least 25 ng/mL had 43% lower risk of developing type 2 diabetes compared to persons with a serum 25(OH)D concentration below 14 ng/mL. Similarly, a cross-sectional analysis of data from a cancer screening trial found that participants with serum 25(OH)D concentrations above 32 ng/mL had half the risk of developing type 2 diabetes, relative to persons with concentrations below 14.8 ng/mL (Brock et al., 2011). A meta-analysis of three cohort studies showed that vitamin D intake greater than 500 IU/day was associated with lower risk of developing diabetes relative to intakes of less than 200 IU/day (Mitri, Muraru and Pittas, 2011).

1.5.4 Bone Health

The role of vitamin D in calcium metabolism and the importance of vitamin D to bone health are well understood. The treatment of rickets with cod liver oil in the early twentieth century eventually led to the discovery of vitamin D (Rajakumar et al., 2007). We now know that rickets in children, and osteomalacia in adults, are the result of a bone mineralization disorder that is directly related to vitamin D deficiency (Maricic, 2008). Though relatively common in the general population, Holick et al., (2005) assert that vitamin D deficiency may be especially common among women receiving treatment for low bone density. Evidence suggests that vitamin D supplementation can reduce the risk of fractures associated with low bone density. One study of older adults found that vitamin D supplementation reduced fracture risk related to osteoporosis by 15% (RR 0.85, 95% CI 0.72 to 0.98) relative to persons not receiving supplementation (Larson, Mosckilde, and Foldspang, 2003).

1.5.5 Kidney Disease

The kidneys are important in vitamin D metabolism because they are responsible for converting a majority of the circulating 25(OH)D into the hormone 1,α,25 dihydroxyvitamin D (Lehman and Meurer, 2010). Thus, it is possible that renal disease could adversely affect the conversion of vitamin D into its hormone metabolite. However, most epidemiological research into the relationship between kidney disease and vitamin D has focused on serum 25(OH)D. Researchers have reported an association between mortality from chronic kidney disease and serum 25(OH)D levels less than 15 ng/mL (Mehrotta et al., 2009). Fiscella, Winters and Ogedegbe (2011) reported an inverse relationship between serum 25(OH)D concentration and albuminuria, a biomarker for decreased kidney function. Their results indicate that persons with a serum vitamin D concentration of less than 20 ng/mL are at least twice as likely to have albuminuria relative to persons with a serum vitamin D concentration of more than 40 ng/mL. They also found that race was no longer a significant predictor for albuminuria after accounting for serum vitamin

D, and suggest that vitamin D deficiency may at least partially explain the observed racial disparities in the prevalence of kidney disease (Fiscella, Winters and Ogedegbe, 2011).

1.6 Data Source

1.6.1 NHANES Background

Data for this study came from the National Health and Nutrition Examination Survey (NHANES), which is administered by the Centers for Disease Control and Prevention (CDC, 2003; CDC, 2005). The purpose of NHANES is to collect nationally representative data about various health and nutrition topics, in addition to certain key demographic and socioeconomic characteristics for survey participants. Our study included data from two cycles of NHANES: 2003-2004 and 2005-2006.

1.6.2 Survey Design

NHANES uses a stratified multistage probability sample of the civilian non-institutionalized population of the United States. Sample selection occurs in four stages. The first stage involves selecting the primary sampling units (PSU) using probability proportional to size (PPS). PSUs are usually individual counties but may be groups of contiguous counties. The second stage involves selection within the PSU. Segments by city block, or equivalent, are selected using PPS. The third stage involves randomly selecting households within each segment. Segments with high proportions of demographic or socioeconomic groups that were targeted for oversampling have higher probabilities of selection. Oversampling means that the higher probabilities of selection are utilized for certain targeted subpopulations. In NHANES the oversampled subpopulations include African Americans, Mexican-Americans, low-income White

Americans, adolescents aged 13-19 and persons age 60 years old and older. The fourth stage involves randomly selecting persons from each household.

1.6.3 Data Collection

Data collection for NHANES took place during a health interview and a medical examination. The health interview was administered to survey participants by a trained survey technician. The interview included a comprehensive questionnaire to gather data on socioeconomic and demographic characteristics, self-identified medical conditions and dietary habits. The medical examination took place in a self-contained mobile examination center (MEC) in order to ensure standardized conditions for data collection. The medical examination consisted of a physical exam, a laboratory component, and various diagnostic testing procedures performed by qualified medical personnel (e.g. physicians, nurses or medical technicians). Additional interview data, such as dietary recall, may have also been collected during the medical examination.

1.7 Rationale for Study

This study addresses several key gaps in the literature related to vitamin D. The contribution of dietary sources to serum vitamin D concentrations is not well studied. This study uses an analysis of dietary patterns to examine the relationship between the consumption of certain types of foods and serum vitamin D concentration. Dietary patterns are also considered as potential risk factors for many medical conditions that have been found to be associated with vitamin D deficiency, including conditions such as diabetes, cancer and cardiovascular disease. Though many studies have addressed the association between vitamin D deficiency and a variety of diseases, we are unaware of any that have controlled for dietary patterns in addition to demographic or behavioral factors. Finally, this study uses an updated release of vitamin D (November 2010) data for the 2003-2004 and 2005-2005 NHANES cycles. Adjustment was made by CDC to account for assay reformulation that took place during these survey cycles.

2 METHODS AND PROCEDURES

2.1 Study Population

The study population was limited to MEC-examined adults aged twenty years old and older who participated in the Food Frequency Questionnaire (FFQ) and who were assigned valid FFQ weights. Pregnant or lactating women were excluded from the analysis due to their special dietary requirements. Persons with missing values in the vitamin D measurements or in several other study variables were also excluded. Table 1 shows the sample size (n) at each step of sequentially applying the selection criteria.

Table 1 Study sample size (raw responses) after sequentially applying selection criteria

Criteria[1]	NHANES Cycle: 2003-2004	NHANES Cycle: 2005-2006	Both Cycles
Age >= 20 years	5,041	4,979	10,020
Valid vitamin D measurement	4,498	4,495	8,993
Valid FFQ survey[2]	3,172	2,858	6,030
Neither pregnant nor lactating	2,994	2,643	5,637
Valid sex and race	2,994	2,643	5,637
Valid BMI measurement	2,942	2,613	5,555
Valid supplement response	2,937	2,613	5,550
Valid season of survey	2,937	2,613	5,550
Valid milk consumption response	2,937	2,613	5,550
Valid physical activity response	2,898	2,577	5,475
Valid smoking response	2,897	2,577	5,474
Valid video screen use response	2,897	2,577	5,474

[1] Valid responses are those with non-missing values
[2] Determined by FFQ weight > 0

2.2 Study Variables

2.2.1 Survey Design Variables

Sample weights should be used to produce unbiased national estimates. Different sets of sample weights are provided by CDC to adjust for unequal probabilities of selection and non-response at various stages, as well as post-stratification to U.S. Census population totals for the survey subdomain. Interview weights and MEC examination weights are two major sets of the survey weights. Additional sets of weights were computed for the fasting sample, the FFQ sample, etc. The Food Frequency Questionnaire contains a large number of food items, and the issue of missing values needs to be properly taken care of. A nonzero weight is assigned to a subject with fewer than ten missing food frequencies, and imputed values are provided to the missing frequencies. In the subsequent dietary analysis of the defined study population, the food frequencies are either actual questionnaire responses or imputed values. Per recommendations in the analytic note from CDC (2006), the weights provided for individual two-year cycles of NHANES were used to calculate a combined four-year weight by dividing the two-year weight in half in order to produce estimates for the U.S. population.

As part of the survey design, the PSUs are stratified according to geographical locations and proportions of the minorities, and two PSUs are selected from each stratum. These design-related units are important design elements and must be accounted for in statistical analysis. In order to protect participant confidentiality, the true design variables are not provided in the public use data sets. The NHANES program has constructed Masked Variance Units (MVU) to approximate the variance estimates obtained by using the true survey design variables. The MVUs consists of the pseudo-stratum variable SDMVSTRA and the pseudo-PSU variable SDMVPSU.

2.2.2 Vitamin D

Data were collected on serum concentration of 25-OH- Vitamin D during the laboratory component of NHANES, during which a blood sample was obtained from participants 1 year old and older. We categorized serum vitamin D concentration according to the classification used by the Institutes of Medicine (2010): a) "severe deficiency" <12 ng/mL, b) "inadequate" 12 to 20 ng/mL, and c) "adequate" > 20 ng/mL. In the multivariate analysis studying effects of dietary patterns on vitamin D concentration, natural log transformed serum vitamin D concentration was used as the response to solve the problem of skewed distribution for the untransformed values.

2.2.3 Diet and Food Frequency

A Food Frequency Questionnaire (FFQ) was used to collect data on dietary intakes of NHNAES study participants over the previous 12 months. These data were publicly released by CDC as two separate data sets. The first data set contains raw survey responses about the intake of various foods over several different time intervals (daily, weekly, monthly, etc.) The second data set contains the output form DietCalc software about the estimated daily dietary intakes for 216 foods based on the raw survey data. The DietCalc software also imputed values for missing responses. Survey participants with more than ten missing values have an FFQ weight of zero and were not considered in this study.

This study uses the estimated daily dietary intakes from DietCalc to develop dietary profiles for use in studying the dietary sources of vitamin D. We categorized the 216 foods in the DietCalc output into thirty major food categories in order to more effectively capture general patterns in food intake. Our categorization was based on the major food categories described by

Ganji, Kafai and McCarthy (2009). We then converted the estimated daily intake to a monthly intake by multiplying the daily intake by a factor of 30.4 (the number of days in an average month.) The Appendix contains a table that shows how the 216 foods from the DietCalc output are assigned to the thirty food major food categories. Two food groups, Creamed Soups and Other Soups, were omitted from the analysis due to their relative unimportance in the dietary profiling analysis.

2.2.4 Demographic Variables

Race/Ethnicity: Respondents were asked to report their racial and ethnic background. This self-reported race and ethnicity was classified into four race/ethnicity categories: Mexican-American, Non-Hispanic Black, Non-Hispanic White, and Other. Non-Hispanic White race/ethnicity was used as the reference category in relative risk analysis.

Gender: Self-reported gender was categorized as Male or Female. Male gender was used as a reference category in relative risk analysis.

Age: Self-reported age at the time of the health interview was categorized into four age groups according to the standard age groups reported by the Institutes of Medicine (2010). These age groups are: a) 20 to 30 years old, b) 31 to 50 years old, c) 51 to 70 years old, and d) older than 70 years of age. The youngest category, 20 to 30 years old, was used as a reference category in relative risk analysis.

Poverty Income Ratio: The Poverty Income Ratio (PIR) is the ratio of family income to the local threshold poverty level income for the family's residence. The PIR was categorized into four levels for the purposes of this analysis: a) $PIR < 1$, b) $1 \leq PIR \leq 2.5$, c) $PIR > 2.5$ and d) "Not Reported". The reference category for relative risk analysis is $PIR > 2.5$. The PIR is the only variable included in the analysis measuring the socioeconomic status.

2.2.5 Behavioral and Health Related Variables

Smoking Status: Data about cigarette smoking behavior were collected during the health interview. Respondents were asked whether they smoked at least 100 cigarettes in their life. Persons who responded “no” were considered to have never smoked. Persons who responded “yes” but who indicated in follow up questions that they did not currently smoke were considered to be former smokers. For current smokers, we calculated pack-years in order to distinguish heavy from light or moderate smokers. We constructed a categorical variable for smoking status including to the following categories: a) never smoked, b) former smoker, c) current smoker less than 20 pack-years, d) current smoker greater than 20 pack-years, and e) current smoker unknown pack-years. The reference category for relative risk analysis was “never smoked.”

Body Mass Index: Data for Body Mass Index (BMI) were collected during the medical examination. BMI is a standard index to measure body adiposity, and is calculated as BMI = Mass (kilograms) / $[\text{height (meters)}]^2$. We collapse BMI into three categories: a) BMI < 25 (normal body weight), b) $25 \leq$ BMI < 30 (overweight), and c) BMI ≥ 30 (obese). The reference category for relative risk analysis was BMI < 25 “normal body weight.”

Physical Activity: During the MEC interview respondents were asked to compare their level of physical activity to others of the same age group. We constructed a dichotomous variable to indicate whether a respondent said that he or she was at least as active as his or her peers.

Computer or Television Screen Viewing: Survey participants were asked to report their approximate number of hours of television and computer use per day, including video games. We categorized screen viewing into the following groups: a) less than 2 hours per day, b) 2 to 3 hours per day, and c) 4 or more hours per day. The category “less than 2 hours per day” was used as the reference category in relative risk analysis.

Milk Consumption: Survey participants were asked how often they consumed milk. We constructed a categorical variable for milk consumption based on the following categories: a) milk consumed daily, b) milk consumed more than once a week but not daily, and c) milk consumed less than once a week. This variable was considered in the descriptive analysis, but was omitted from the multivariate analysis because milk consumption was part of the dietary patterns.

Season of Survey: The season during which data were collected for a respondent is reported in NHANES as either fall/winter (October through March) or spring/summer (April through September). The season of data collection is an important control variable due to increased sun exposure during the spring and summer months for many people. Spring/summer was used as the reference category in relative risk analysis.

2.2.6 Medical Conditions

During the health interview, survey respondents were asked whether they currently have or had ever experienced certain medical conditions. We constructed dichotomous variables based on these responses to indicate if a respondent had experienced a particular condition. The dichotomous variables were created for the following medical conditions: cancer, cardiovascular disease, diabetes, kidney disease, or osteoporosis. Persons were considered to have reported cardiovascular disease if they reported any of the following conditions: angina, congestive heart failure, coronary artery disease, or heart attack.

2.3 Statistical Methods

2.3.1 Overview

SAS version 9.2 was used to conduct the statistical analysis for this research. The stratified multistage probability sample design of NHANES means that conventional statistical techniques were not appropriate for these data because these techniques are applied under the assumption of simple random sampling. Procedures for the analysis of complex survey data are available in SAS. These procedures use information about the stratum, primary sampling unit and survey weight to account for the survey design and unequal probabilities of selection in calculating interval estimation and conducting hypothesis tests. The standard Taylor series linearization method was used to find approximate variance estimates.

2.3.2 Descriptive Analysis

The descriptive analysis included analyzing population frequencies and proportions for significant interaction between serum vitamin D concentration and independent variables. This analysis used the Rao-Scott chi square test (Rao and Scott, 1979; Lohr, 1999) to identify significant dependent relationships between two categorical variables. The Rao-Scott chi square test accounts for the complex survey design, and the test statistic (χ^2_{RS}) is calculated as

$$X^2_{RS} = \frac{(r-1)(c-1)X^2}{E[X^2]}$$

where r is the number of rows and c is the number of columns of the two-way table, and X^2 is the conventional chi square statistic under the assumption of simple random sampling. E[X^2] is the expected value of X^2 given the complex survey design, and is calculated as

$$E[X^2] \approx \sum_{i=1}^{r}\sum_{j=1}^{c}(1-p_{ij})d_{ij} - \sum_{i=1}^{r}(1-p_{i+})d_i^R - \sum_{j=1}^{c}(1-p_{+j})d_j^C$$

where d_{ij} is the design effect for proportion p_{ij}, d_i^R is the design effect for proportion p_{i+}, and d_j^C is the design effect for proportion p_{+j}. The design effect can be obtained for different statistics and varies by study design. For the Rao-Scott chi square statistic, the design effect is the ratio between the variance under the given design and the variance based on a simple random sample. If the ratio is evaluated less than one, the study design yields more precise estimation result than a simple random sample. Otherwise, the study design is less efficient in estimation than a simple random sample. Under the null hypothesis the Rao-Scott chi square will approximately follow a chi square distribution with $(r - 1)(c - 1)$ degrees of freedom. The SAS implementation of the Rao-Scott chi square test is available in the SURVEYFREQ procedure.

2.3.3 Principal Components Analysis

In our analysis we sought to determine general dietary patterns based on estimated monthly dietary intakes for twenty-eight major food categories. We used Principal Components Analysis (PCA) to summarize the information contained in these twenty-eight food categories into dietary patterns that could be used to classify NHANES participants by the types and frequency of the foods they consumed. By using PCA we also mitigate the bias introduced by mulicollinearity associated with foods typically consumed together during meals or for other reasons. Our approach to constructing dietary profiles is similar to that used by Ganji, Kafai and McCarthy (2009) and Cho, Chin and Kim (2011). Both approaches use PCA to determine dietary patterns from food frequency data. However, Ganji, Kafai and McCarthy (2009) determine a dominant pattern for each respondent based on their component scores. Our approach is more similar to Cho, Chin and Kim (2011), who do not determine a dominant pattern but instead consider the relative contribution of each pattern to the respondent's overall diet. In addition to determining component scores, we also categorize each component by tertile for low, medium and high scores for use as a categorical variable.

Principal components analysis is an often-used multivariate dimension reduction technique that is available in many statistical software packages. The basic idea behind PCA is to derive one or more important factors that capture a substantial portion of the variation among a group of variables. The factors, or principal components, are uncorrelated linear combinations of variables. We used the FACTOR procedure implementation of PCA in SAS to conduct the analysis for this study. Assuming a matrix $\mathbf{X}$ of p variables that are standardized to have a mean 0 and unit variance, the factor loadings for the principal components are the eigenvectors of the correlation matrix $\mathbf{X}^{\mathbf{T}}\mathbf{X}$, and the variance of each principal component corresponds to the eigenvalues of $\mathbf{X}^{\mathbf{T}}\mathbf{X}$ (Montgomery, Peck and Vining, 2006). The eigenvalues ($\lambda_1, \lambda_2, \ldots, \lambda_p$), are ordered in descending order by magnitude. Typically only a subset of the principal components is retained, and the number of principal components retained for use in an analysis is determined by decision criteria based on the variation explained by the subset of eigenvalues or on the size of individual eigenvalues. In our analysis we retained all principal components that explained at least 5% of the variance. We applied an orthogonal varimax rotation to the retained factors to make them easier to interpret.

2.3.4 Regression Modeling

Multiple regression was used to examine the relationship between the demographic, behavioral and dietary factors and natural log transformed serum vitamin D concentration. For a response Y the multiple regression model for k predictors $x_1, x_2, \ldots, x_k$ is (Montgomery, Peck and Vining, 2006):

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

In this model the β_0 term represents the intercept, the $\beta_1 \dots \beta_k$ terms are the regression coefficients, and ε is an error term. The survey design-adjusted implementation of multiple regression in SAS is the procedure SURVEYREG. The NHANES design uses multi-stage sampling. Suppose that there are H strata, n_h PSUs and m_{hi} subjects within the ith PSU of the hth stratum. and PSU. The design-based estimate for the vector of the regression coefficients, **B,** is given by:

$$\widehat{\mathbf{B}} = \left(\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}\, \mathbf{x}_{hij} \mathbf{x}_{hij}^{T} \right)^{-1} \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}\, \mathbf{x}_{hij} y_{hij}$$

where w_{hij} is the sampling weight, $\mathbf{x}_{hij}$ is the vector of predictor variables, and y_{hij} is the response variable (Lohr, 1999). Assume a vector $\mathbf{q}_{hij}$ defined as:

$$\mathbf{q}_{hij} = \mathbf{x}_{hij}\left(y_{hij} - \mathbf{x}_{hij}^{T} \widehat{\mathbf{B}}\right).$$

The covariance matrix of $\widehat{\mathbf{B}}$ can be estimated by $\hat{V}\left(\widehat{\mathbf{B}}\right)$,

$$\hat{V}\left(\widehat{\mathbf{B}}\right) = \left(\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}\, \mathbf{x}_{hij} \mathbf{x}_{hij}^{T} \right)^{-1} \hat{V} \left(\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}\, \mathbf{q}_{hij} \right) \left(\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}\, \mathbf{x}_{hij} \mathbf{x}_{hij}^{T} \right)^{-1}$$

(Lohr, 1999).

Logistic regression was used to model the association between serum vitamin D concentration and self-reported medical condition. The survey design-adjusted implementation of logistic regression in SAS is the procedure SURVEYLOGISTIC. This procedure provides design-corrected estimates of odds ratios and confidence intervals. Assuming a dichotomous response Y, the logistic model for k predictors $x_1, x_2, \dots, x_k$ is (Agresti, 2007):

$$log\left(\frac{P(Y=1)}{P(Y=0)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

In this model the probability of an event is P(Y=1), and the response is the natural logarithm of the odds of the event. The β_0 term represents the intercept, the $\beta_1 \dots \beta_k$ terms are the regression coefficients, and ε is an error term. The survey design-based estimates of regression

cients $(\widehat{\mathbf{B}})$ are obtained by solving the estimating equation,

$$\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij}\, x_{hijk}\left[y_{hij} - \frac{\exp(\mathbf{x}_{hij}^{T}\widehat{\mathbf{B}})}{1+\exp(\mathbf{x}_{hij}^{T}\widehat{\mathbf{B}})}\right] = 0$$

for k = 1,...,p predictors. We refer the reader to Lohr (1999) for the survey design-adjusted variance estimation.

3 RESULTS

3.1 Descriptive Analysis

3.1.1 Characteristics of Study Population

A total of 5,474 respondents met the criteria for inclusion in the study. Table 2 shows the demographic, behavioral, environmental profiles for study participants. By weighted percentage, females represented a slight majority of respondents (53.1%), while approximately 38.4% of respondents belonged to the 31 to 51 year old age group. Non-Hispanic Whites were the most common race/ethnicity at 73.9%, followed by non-Hispanic Blacks (11.0%), Mexican Americans (7.5%) and Other race/ethnicity (7.5%). Body Mass Index (BMI) was evenly distributed across the three categories corresponding to "normal" (BMI < 25), "overweight" (BMI 25 to 30) and "obese" (BMI > 30). Most participants (57.3%) were from households with a poverty income ratio above 2.5, while poverty income ratios of less than 1, indicating a household income at or below the local poverty level, represented 10.6% of respondents. More than three quarters of respondents (78.3%) indicated that they were at least as physically active as other persons of the same age. More than half of respondents (56.0%) reported taking a dietary supplement. A slight majority of respondents (52.5%) reported more than 3 hours of video screen viewing per day. Fewer than half of respondents reported daily milk consumption, while more than three quarters (75.4%) indicated that they were not current smokers. About sixty percent of respondents were interviewed and examined during the spring and summer months, when ultraviolet light from sun exposure is most intense. Most respondents (62.7%) were found to have serum vitamin D levels that were above 20 ng/mL, while those with inadequacy (serum vitamin

Table 2 Characteristics of study participants

Variable	Value	Frequency	Weighted Percent	Weighted Frequency
Gender	Male	2,705	46.9	89,728,506
	Female	2,769	53.1	101,396,599
Age group	20-30 years	892	19.7	37,602,613
	31-50 years	1,830	38.4	73,421,146
	51-70 years	1,702	31.0	59,211,233
	>70 years	1,050	10.9	20,890,113
Race/ethnicity	Non-Hispanic White	3,133	73.9	141,278,658
	Non-Hispanic Black	1,047	11.0	21,101,463
	Mexican-American	933	7.5	14,356,519
	Other	361	7.5	14,388,465
BMI	<25	1,700	33.6	64,297,816
	25 to <30	1,884	33.1	63,356,394
	30+	1,890	33.2	63,470,895
Poverty income ratio	Not Reported	236	4.1	7,814,440
	PIR <1	788	10.6	20,283,113
	PIR 1-2.5	1,810	28.0	53,523,542
	PIR >2.5	2,640	57.3	109,504,010
At least as physically active as peers	No	1,150	21.7	41,466,062
	Yes	4,324	78.3	149,659,043
Any supplement use	No	2,525	44.0	84,028,317
	Yes	2,949	56.0	107,096,788
Hours of video or computer screen use per day	2 or fewer hours	2,546	47.5	90,827,738
	3-4 hours	1,693	31.9	60,965,257
	4+ hours	1,235	20.6	39,332,110
Milk consumption	Less than once a week or varied	1,574	28.3	54,102,405
	At least once a week but not daily	1,455	27.7	53,009,331
	Milk consumed daily	2,445	44.0	84,013,369
Smoking	Never	2,737	49.4	94,346,556
	Former	1,523	26.0	49,708,833
	Current <20 pack-years	141	2.7	5,239,220
	Current 20+ pack-years	845	17.7	33,771,963
	Current unknown pack-years	228	4.2	8,058,533
Six month time period	Fall/Winter	2,444	40.6	77,679,466
	Spring/Summer	3,030	59.4	113,445,639
Serum vitamin D Concentration	<12 ng/mL	685	8.7	16,631,215
	12-20 ng/mL	1,810	28.6	54,630,890
	>20 ng/mL	2,979	62.7	119,863,000

D between 12 and 20 ng/mL) or severe deficiency (serum vitamin D less than 12 ng/mL) represented 28.6% and 8.7% of the population, respectively.

3.1.2 Vitamin D and Demographic, Behavioral and Environmental Variables

Table 3 shows the descriptive analysis of the relationship between serum vitamin D concentration and the demographic, behavioral and environmental variables. A significant relationship was found between serum vitamin D concentration and gender ($p<0.0001$), race/ethnicity ($p<0.0001$), BMI ($p<0.0001$), poverty income ratio ($p<0.0001$), physical activity ($p<0.0001$), supplement use ($p<0.0001$), video screen viewing ($p<0.0001$), milk consumption ($p<0.0001$), cigarette smoking ($p=0.0001$), and season of survey administration ($p<0.0001$). Age group was not found to be significantly associated with serum vitamin D concentration.

3.1.3 Vitamin D and Food Groups

Table 4 shows the relationship between serum vitamin D concentration and mean monthly servings of twenty-eight food groups consumed by survey respondents. The foods that were significantly associated with serum vitamin D concentration include: Alcohol ($p<0.0001$), Cereals ($p<0.0001$), Coffee or Tea ($p<0.0001$), Condiments ($p=0.0003$), Cruciferous and Green Vegetables ($p=0.0018$), Dairy Alternative Meal Replacement ($p=0.0042$), Eggs ($p=0.0043$), Energy Drinks ($p<0.0001$), Fruit ($p=0.0147$), Fruit Juice ($p<0.0001$), High Far Dairy ($p=0.0002$), Legumes ($p<0.0001$), Low Fat Dairy ($p<0.0001$), Meat ($p<0.0001$), Mixed Foods ($p<0.0001$), Nuts ($p<0.0001$), Poultry ($p=0.0011$), Processed Meats ($p<0.0001$), Refined Grains ($p<0.0001$), Tomatoes ($p=0.0114$) and Whole Grains ($p<0.0001$). Foods that were not significantly associated with serum vitamin D concentration include: Butter or Margarine ($p=.3013$), Fish and Other Seafood ($p=0.3847$), Other Fats ($p=0.0822$), Other Vegetables ($p=0.5128$), Pizza ($p=0.06$), Snacks and Sweets ($p=0.6576$) and Starchy Vegetables ($p=0.0866$).

Table 3 Weighted percentages of respondents by demographic and behavioral characteristics within serum vitamin D concentration category

Variable	Value	Serum Vitamin D Concentration			P_{Trend}[1]
		<12 ng/mL	12-20 ng/mL	>20 ng/mL	
Gender	Male	6.8	27.7	65.5	<.0001
	Female	10.4	29.4	60.3	
Age group	20-30 years	8.9	30.4	60.6	0.6686
	31-50 years	9.4	28.4	62.2	
	51-70 years	7.9	27.8	64.3	
	>70 years	8.1	28.2	63.7	
Race/ethnicity	Non-Hispanic White	3.8	23.3	72.9	<.0001
	Non-Hispanic Black	36.6	44.1	19.4	
	Mexican-American	14.7	42.8	42.4	
	Other	10.3	43.5	46.3	
BMI	<25	5.6	23.4	71.0	<.0001
	25 to <30	5.7	26.3	68.0	
	30+	14.9	36.1	49.0	
Poverty income ratio	Not Reported	8.6	33.2	58.2	<.0001
	PIR <1	13.9	36.5	49.6	
	PIR 1-2.5	11.7	29.9	58.4	
	PIR >2.5	6.3	26.2	67.6	
At least as physically active as peers	No	14.5	37.1	48.4	<.0001
	Yes	7.1	26.2	66.7	
Any supplement use	No	14.2	34.4	51.4	<.0001
	Yes	4.4	24.0	71.6	
Hours of video or computer screen use per day	2 or fewer hours	6.8	25.1	68.2	<.0001
	3-4 hours	7.9	30.7	61.5	
	4+ hours	14.5	33.5	52.1	
Milk consumption	Less than once a week or varied	14.6	31.6	53.8	<.0001
	At least once a week but not daily	9.4	31.3	59.3	
	Milk consumed daily	4.5	24.9	70.6	
Smoking	Never	8.6	28.6	62.8	0.0001
	Former	6.8	25.7	67.5	
	Current <20 pack-years	8.2	27.3	64.5	
	Current 20+ pack-years	11.8	34.0	54.2	
	Current unknown pack-years	8.7	24.3	66.9	
Six month time period	Fall/Winter	12.1	36.3	51.6	<.0001
	Spring/Summer	6.4	23.3	70.4	

[1] Significance of Rao-Scott chi square test

Table 4 Mean food group servings per month by serum vitamin D level

Food Group	Serum Vitamin D Concentration			P_{Trend}[1]
	<12 ng/mL	12-20 ng/mL	>20 ng/mL	
Alcohol	10.5	8.7	12.9	<.0001
Butter or Margarine	24.9	23.4	23.8	0.3013
Cereals	9.7	12	13.2	<.0001
Coffee or Tea	50.8	59	64.3	<.0001
Condiments	37	38.3	34.2	0.0003
Cruciferous or Green Vegetables	19.6	20.8	21.9	0.0018
Dairy Alternative Meal Replacement	12.2	16	15.7	0.0042
Eggs	6.1	7	6.7	0.0043
Energy Drinks	51.2	46.1	35.4	<.0001
Fish and Other Seafood	5.8	5.5	5.4	0.3847
Fruit	40.6	39.5	42.3	0.0147
Fruit Juice	27.8	23.1	20	<.0001
High Fat Dairy	28.3	33.8	34.3	0.0002
Legumes	4.2	4.2	3.4	<.0001
Low Fat Dairy	17.1	29.5	37.8	<.0001
Meat	15.1	14.5	13.4	<.0001
Mixed Foods	12.9	13.2	11.8	<.0001
Nuts	6.7	7.7	9.8	<.0001
Other Fats	24.9	25.8	24.9	0.0822
Other Vegetables	40.1	39.5	38.9	0.5128
Pizza	2.1	2.1	2.3	0.06
Poultry	7.8	7.1	6.7	0.0011
Processed Meat	18	16.7	14.7	<.0001
Refined Grains	29.8	27.1	26	<.0001
Snacks or Sweets	33.5	32.7	32.6	0.6576
Starchy Vegetables	24.4	23.5	24.4	0.0866
Tomatoes	19.1	21.3	21.1	0.0114
Whole Grains	9.5	11.6	12.6	<.0001

[1] Significance of Rao-Scott chi square test

3.2 Dietary Profiling Using Principal Components Analysis

3.2.1 Principal Components Analysis

Foods are often consumed together as parts of meals or for other reasons, and the consumption of certain foods may be highly correlated with others. This high degree of correlation among foods means that multicollinearity is a concern when analyzing dietary data for individual foods or food groups. Rather than looking at single foods in isolation, we used principal components analysis to determine patterns of food consumption based on the data from the Food Frequency Questionnaire. Table 5 shows the first ten principal components extracted from our data set of twenty-eight food group variables. The first three principal components capture a substantial portion of the variation in the food frequency data and additional components capture an incrementally smaller amount. In fact, the first three principal components represent more than one-third of the variation in the dietary data.

Table 5 Summary of first ten principal components

Principal Component[1]	Eigenvalue	Proportion	Cumulative Proportion
1	5.97	0.21	0.21
2	2.01	0.07	0.29
3	1.87	0.07	0.35
4	1.41	0.05	0.40
5	1.26	0.05	0.45
6	1.08	0.04	0.49
7	1.05	0.04	0.52
8	0.96	0.03	0.56
9	0.92	0.03	0.59
10	0.88	0.03	0.62

[1] There was a total of 28 principal components

Since they each described at least 5% of the total variation in the data, we retained the first three principal components for use in profiling the dietary patterns of survey respondents. We applied a varimax rotation to the retained principal components in order to improve their interpretability. After rotation, Factor 1 accounted for about 14.0% of the variation, and Factors 2 and 3 accounted for 13.3% and 7.8% respectively. The factor pattern for the transformed components appears in Table 6. We considered each factor to contain a characteristic mix of foods that could be interpreted as a distinct dietary pattern, and that the contribution of the three patterns taken together could describe the overall dietary pattern for each respondent. Foods with higher factor loadings contribute the most to the overall score for a factor and are shown in bold. Negative scores can be interpreted as foods that are negatively associated with a pattern. In other words, the presence of these foods in a person's diet could be interpreted as meaning that their diet is less similar to a particular pattern. We can very generally summarize the underlying dietary pattern of the factor based on which foods have the highest scores. Factor 1 is characterized by fruits, vegetables and lean meats, including the following foods with scores higher than 0.5: Cruciferous and Green Vegetables, Fruit, and Other Vegetables. Factor 2 is characterized by high fat meats and processed foods, including the following foods with scores higher than 0.5: Energy Drinks, Meat, Mixed Foods, Pizza, Processed Meats, Refined Grains, Snacks and Sweets, and Starchy Vegetables. Factor 3 mainly describes beverages, condiments and food additives, including the following foods with scores of 0.5 or higher: Coffee and Tea, Condiments, and Dairy Alternative Meal Replacements. Since these factors represent patterns of food consumption we refer to them as "dietary patterns."

Factor scores were used to construct a categorical variable for each pattern based on tertiles for the dietary pattern score, corresponding to high, medium and low scores. The categorical dietary pattern variables were used in subsequent statistical analysis.

Table 6 Factor pattern for dietary profile principal components after varimax rotation

Food Group	Factor 1	Factor 2	Factor 3
Alcohol	-0.03	0.20	0.00
Butter or Margarine	0.13	0.28	0.43
Cereals	0.44	-0.04	0.03
Coffee or Tea	0.05	-0.08	**0.73***
Condiments	0.02	0.03	**0.77***
Cruciferous or Green Vegetables	**0.73***	0.04	0.04
Dairy Alternative Meal Replacement	0.02	-0.09	**0.55***
Eggs	0.19	0.17	0.29
Energy Drinks	-0.17	**0.52***	0.00
Fish and Other Seafood	0.43	0.36	-0.02
Fruit	**0.73***	0.12	0.00
Fruit Juice	0.36	0.32	-0.08
High Fat Dairy	-0.02	0.28	0.47
Legumes	0.42	0.09	-0.04
Low Fat Dairy	0.38	-0.07	0.09
Meat	0.14	**0.75***	0.12
Mixed Foods	0.41	**0.62***	0.05
Nuts	0.46	0.04	0.18
Other Fats	0.41	0.30	0.28
Other Vegetables	**0.69***	0.17	0.15
Pizza	-0.02	**0.53***	-0.06
Poultry	0.38	0.45	-0.02
Processed Meat	0.14	**0.64***	0.15
Refined Grains	0.19	**0.58***	0.21
Snacks or Sweets	0.25	**0.54***	0.14
Starchy Vegetables	0.47	**0.52***	0.12
Tomatoes	0.44	0.29	0.06
Whole Grains	0.47	0.02	0.04

* Factor loading > 0.5 indicates important food group for dietary pattern

3.2.2 Descriptive Analysis of Vitamin D and Dietary Patterns

Table 7 shows the weighted percentage of respondents, by vitamin D status, within each dietary pattern score category. The trend across score categories in all three dietary patterns was significantly associated with serum vitamin D. Higher scores for patterns 1 and 3 were associated with

Table 7 Weighted percentages of respondents by dietary pattern score within serum vitamin D level

Variable	Value	Serum Vitamin D Concentration			P_{Trend}[1]
		<12 ng/mL	12-20 ng/mL	>20 ng/mL	
Dietary Pattern 1	Low	12.3	30.8	56.9	<.0001
	Medium	7.1	28.4	64.6	
	High	6.7	26.6	66.7	
Dietary Pattern 2	Low	6.6	26.1	67.3	<.0001
	Medium	8.3	30.0	61.8	
	High	11.2	29.7	59.2	
Dietary Pattern 3	Low	11.7	29.6	58.7	<.0001
	Medium	6.8	28.9	64.3	
	High	7.7	27.3	65.1	

[1] Significance of Rao-Scott chi square test

higher vitamin D concentrations, while higher scores for pattern 2 were associated with lower vitamin D concentrations.

3.3 Regression Models for Serum Vitamin D Concentration

3.3.1 Demographic, Behavioral and Dietary Determinants of Vitamin D

We used multivariable regression analysis to study the relationship between serum vitamin D concentration and the demographic, behavioral, environmental and dietary variables. The response variable for this analysis was the natural logarithm of the serum 25(OH)-Vitamin D concentration for each respondent. Table 8 shows the regression coefficients for the fitted model that included significant main effects only. The multivariate analysis revealed that serum 25(OH)-Vitamin D concentration was associated with gender, age group, race/ethnicity, BMI, physical activity, supplement use, video screen viewing, seasonality, Dietary Pattern 1 (fruits, vegetables and lean meat) and Dietary Pattern 2 (high fat meats and processed foods). Statistical significance was assessed by a t test for the individual regression coefficient.

Table 8 Estimated regression model coefficients for model to predict the natural log transformed serum vitamin D concentration

Parameter[1]	Value	Estimate[2]	Standard Error	P value[3]
Intercept		3.28	0.03	<.0001
Gender	Male (Reference)			
	Female	-0.05	0.01	<.0001
Age group	20-30 years (Reference)			
	31-50 years	-0.02	0.02	0.3851
	51-70 years	-0.07	0.02	0.0068
	>70 years	-0.13	0.03	<.0001
Race/ethnicity	Non-Hispanic White (Reference)			
	Non-Hispanic Black	-0.5	0.03	<.0001
	Mexican-American	-0.22	0.02	<.0001
	Other	-0.24	0.03	<.0001
BMI	<25 (Reference)			
	25 to <30	-0.04	0.01	0.0003
	30+	-0.17	0.01	<.0001
At least as physically active as peers	Yes (Reference)			
	No	-0.13	0.02	<.0001
Any supplement use	No (Reference)			
	Yes	0.12	0.01	<.0001
Hours of video or computer screen use per day	4+ hours (Reference)			
	3-4 hours	0.02	0.02	0.2529
	2 or fewer hours	0.05	0.02	0.0047
Six month time period	Spring/Summer (Reference)			
	Fall/Winter	-0.12	0.03	<.0001
Dietary Pattern 1 (fruits, vegetables and lean meats)	Low (Reference)			
	Medium	0.05	0.02	0.0039
	High	0.04	0.02	0.012
Dietary Pattern 2 (high fat meats and pro-cessed foods)	Low (Reference)			
	Medium	-0.04	0.01	0.0084
	High	-0.06	0.02	0.0061
Dietary Pattern 3 (condiments and food additives)	Low (Reference)			
	Medium	0.03	0.02	0.076
	High	0.03	0.01	0.0552

[1] Model includes significant predictors only. Poverty income ratio and smoking were not significantly associated with serum vitamin D concentration and were not included in the final model. Milk consumption was not considered as a separate predictor due to the inclusion of high and low fat dairy consumption in the dietary patterns.
[2] Natural logarithmic scale
[3] Observed significance based on t test

Relative to reference categories, an inverse relationship with serum vitamin D concentration was found for non-Hispanic Black race/ethnicity ($p<0.0001$), Mexican American race/ethnicity ($p<0.0001$), Other race/ethnicity ($p<0.0001$), obesity or BMI between 25 and 30 ($p=0.0003$), overweight or BMI of at least 30 ($p<0.0001$), lower levels of physical activity ($p<0.0001$), Fall/Winter seasonality ($p<0.0001$), and higher scores for Dietary Pattern 2 characterized by high fat meats and processed foods ($p=0.0061$). Younger age ($p<0.0001$), supplement use ($p<0.0001$), fewer hours of video screen viewing ($p=0.0047$) and higher scores for Dietary Pattern 1 characterized by fruits, vegetables and lean meat ($p=0.0039$) were associated with higher concentrations of serum vitamin D.

3.3.2 Dietary Patterns and Vitamin D

To assess the relationship between dietary pattern and vitamin D status we looked at the trend in geometric mean serum vitamin D concentration across categories representing tertiles (low, medium and high) for dietary pattern scores. Table 9 shows the unadjusted geometric means (for comparison) and table 10 shows the model-adjusted geometric mean serum vitamin D concentration within each dietary pattern. Unadjusted geometric means for each pattern all exhibited significant trends. Higher scores for Dietary Pattern 1 (fruits, vegetables and lean meat) and Dietary Pattern 3 (beverages, condiments and food additives) were associated with higher geometric mean serum vitamin D concentrations ($p=0.0005$ and $p=0.0013$, respectively). There was an inverse relationship between scores for Dietary Pattern 2 (high fat meats and processed foods) and geometric mean serum vitamin D concentration ($p<0.0001$). After Bonferroni adjustment for multiple comparisons, there was a significant difference between means for low levels of both Dietary Patterns 1 and 3 versus medium and high levels, but not between medium and high. This result suggests that scores above the middle tertile for either Dietary Pattern 1 or 3 have little additional influence on vitamin D status. Means across low, medium and high levels for Dietary Pattern 2 were each significantly different from one another, which suggests an incremental effect on vitamin D that increases as the pattern score increases.

The model used for multivariate adjustment included age, gender, race/ethnicity, BMI, video screen viewing, supplement use, physical activity, seasonality, and the other dietary patterns. The model included interactions for age and gender, age and supplement use, gender and BMI, gender and supplement use, and video screen viewing and supplement use. After multivariate adjustment Dietary Pattern 1 was positively associated with mean serum vitamin D concentration (p=0.0063), while Dietary Pattern 2 was negatively associated with Vitamin D status (p=0.0166). Dietary Pattern 3 was not significantly associated with vitamin D after multivariate adjustment. After Bonferroni adjustment for multiple comparisons, there was a significant difference between means for low levels of both Dietary Patterns1 and 2 versus medium and high levels, but not between medium and high. This result suggests that scores above the middle tertile for either Dietary Pattern have little additional influence on vitamin D status.

Table 9 Unadjusted geometric mean serum vitamin D concentration by dietary pattern score

	Dietary Pattern	**Dietary Pattern Score**			$\mathbf{P_{Trend}}$[1]
		Low	**Medium**	**High**	
Unadjusted mean serum vitamin D concentration (± SE)	Dietary Pattern 1(Fruits, Vegetables and Lean Meats)	$20.7^{a,b}$ ± 0.6	22.6^{a} ± 0.4	22.8^{b} ± 0.4	0.0005
	Dietary Pattern 2 (High Fat Meats and Processed Foods)	$23.3^{a,b}$ ± 0.5	$21.8^{a,c}$ ± 0.4	$20.9^{b,c}$ ± 0.5	<.0001
	Dietary Pattern 3 (Beverages, Condiments and Food Additives)	$20.9^{a,b}$ ± 0.6	22.7^{a} ± 0.4	22.4^{b} ± 0.4	0.0013

[1] Significance of F ratio for effect in model
[a, b, c] Means with the same subscript are significantly different at p≤0.05 after Bonferroni adjustment for multiple comparisons

Table 10 Multivariate adjusted geometric mean serum vitamin D concentration by dietary pattern score

	Dietary Pattern	Dietary Pattern Score			P_{Trend}[1]
		Low	Medium	High	
Adjusted[2] mean serum vitamin D concentration (± SE)	Dietary Pattern 1 (Fruits, Vegetables and Lean Meats)	$16.9^{a,b}$ ± 0.3	17.9^{a} ± 0.3	17.8^{b} ± 0.3	0.0063
	Dietary Pattern 2 (High Fat Meats and Processed Foods)	$18.1^{a,b}$ ± 0.3	17.4^{b} ± 0.3	17.1^{a} ± 0.4	0.0166
	Dietary Pattern 3 (Beverages, Condiments and Food Additives)	17.2 ± 0.3	17.7 ± 0.3	17.6 ± 0.3	0.1488

[1] Significance of F ratio for effect in model
[2] Adjusted for age, gender, race/ethnicity, BMI, video screen viewing, supplement use, physical activity, seasonality, and the other dietary patterns. There were significant interaction terms for age and gender, age and supplement use, gender and BMI, gender and supplement use, and video screen viewing and supplement use.
[a, b] Means with the same subscript are significantly different at p≤0.05 after Bonferroni adjustment for multiple comparisons

3.4 Vitamin D and Self-Reported Medical Conditions

3.4.1 Descriptive Analysis

The sample size, raw frequency, percentage prevalence (weighted percentage) and weighted frequency for the five self-reported medical conditions in our study are displayed in Table 11. There were 530 respondents (n = 5,466) who reported ever having been clinically diagnosed with any type of cancer. These cases represented an overall cancer prevalence of about 8.6% in the general population. There were 579 respondents (n = 5,473) who reported having been told by a doctor that they had cardiovascular disease, which represented a prevalence of 7.8% in the general population. There were 567 respondents (n = 5,474) who reported having been told by a doctor that they had diabetes, corresponding to a prevalence for diabetes of 7.5% in the general population. Kidney disease was relatively uncommon and was reported by 144 respondents (n = 5,474) for a population prevalence of 2.0%. Physician diagnosed Osteoporosis was reported by 380 respondents (n = 5,474) for a population prevalence of 5.7%.

Table 11 Sample sizes and cases for self-reported medical conditions

Medical Condition	Sample Size	Cases	Weighted Percent	Weighted Frequency
Cancer	5,466	530	8.6	16,435,206
Cardiovascular disease	5,473	579	7.8	14,869,696
Diabetes	5,474	567	7.5	14,357,194
Kidney disease	5,474	144	2.0	3,785,013
Osteoporosis	5,474	380	5.7	10,910,395

Table 12 shows the weighted percentage of respondents within each of the three categories of vitamin D sufficiency who reporting one of the five medical conditions. A Rao-Scott chi square test was used to assess the significance of the trend across vitamin D sufficiency categories in the proportion of respondents who reported a medical condition, and the p value for trend is reported in the rightmost column of the table. Self-reported medical conditions that were significantly associated with serum vitamin D concentration include: cancer ($p=0.0025$), cardiovascular disease ($p=0.0004$), diabetes ($p<0.0001$), and kidney disease ($p=0.0014$). Osteoporosis was not significantly associated with serum vitamin D concentration ($p=0.6552$). For cardiovascular disease, diabetes and kidney disease lower serum vitamin D concentrations were associated with a larger proportion of respondents reporting the medical condition. Conversely, higher serum vitamin D concentrations were associated with increased cancer prevalence.

3.4.2 Logistic Regression Analysis of Vitamin D and Self-Reported Medical Conditions

In the descriptive analysis we saw that many of the self-reported medical conditions in the NHANES survey were associated with lower serum vitamin D levels. However, since many factors that are associated with serum vitamin D concentration are also associated with many chronic diseases and other medical conditions, adjustment for the effects of confounding variables is necessary. We used logistic regression to study the relationship between self-reported

Table 12 Weighted percentages of respondents reporting medical conditions within serum vitamin D level

	Serum Vitamin D Concentration			
Medical Condition	**<12 ng/mL**	**12-20 ng/mL**	**>20 ng/mL**	P_{Trend}[1]
Cancer	4.9	7.7	9.6	0.0025
Cardiovascular disease	11.7	9.2	6.6	0.0004
Diabetes	11.4	9.9	5.9	<.0001
Kidney disease	3.1	2.8	1.4	0.0014
Osteoporosis	5.1	5.4	5.9	0.6552

[1] Significance of Rao-Scott chi square test

medical condition (outcome) and vitamin D (predictor) after controlling for these potential confounders. A separate logistic regression model was constructed for each medical condition. Candidate variables for logistic regression analysis included age, gender, race/ethnicity, poverty income ratio, BMI, smoking, video screen viewing, supplement use, physical activity, and the three dietary patterns. Any significant interactions between these variables were also considered in model construction. The significance of trending in the severity of vitamin D deficiency was assessed by using the Wald chi square statistic.

Table 13 shows the unadjusted and adjusted odds ratios for the association between self-reported cancer and serum vitamin D concentration. Before adjustment for confounding variables, respondents with a serum vitamin D concentration between 12 and 20 ng/mL were not significantly more likely to report cancer (OR 0.79, 95% CI 0.61 to 1.03) relative to those with adequate levels of vitamin D. Respondents with a serum vitamin D concentration less than 12 ng/mL were significantly less likely to report cancer (OR 0.49, 95% CI 0.32 to 0.73) relative to those with adequate levels of vitamin D before multivariate adjustment. However, after multivariate adjustment there was no longer a significant association between vitamin D deficiency and self-reported cancer. The multivariate logistic regression model included adjustment for age,

Table 13 Unadjusted and adjusted odds ratios for the association between serum vitamin D concentration and self-reported cancer

Cancer		Odds Ratio	95% Wald Confidence Limits		P_{Trend}[1]
			Lower	Upper	
Unadjusted	Vitamin D >20 ng/mL	(Reference)	-	-	0.0025
	Vitamin D 12-20 ng/mL	0.789	0.605	1.029	
	Vitamin D <12 ng/mL	0.486	0.323	0.732	
Adjusted[2]	Vitamin D >20 ng/mL	(Reference)	-	-	0.1489
	Vitamin D 12-20 ng/mL	0.898	0.663	1.215	
	Vitamin D <12 ng/mL	0.601	0.358	1.008	

[1] Significance of Wald chi square statistic
[2] Adjusted for age, gender, race/ethnicity, supplement use, and smoking. The model included significant interactions for age and race, age and supplement use, age and smoking, gender and race and race and smoking.

gender, race/ethnicity, supplement use, and smoking. The model included interactions for age and race, age and supplement use, age and smoking, gender and race and race and smoking.

Table 14 shows the unadjusted and adjusted odds ratios for the association between self-reported cardiovascular disease and serum vitamin D concentration. Before adjustment for confounding variables, respondents with a serum vitamin D concentration between 12 and 20 ng/mL were significantly more likely to report cardiovascular disease (OR 1.45, 95% CI 1.13 to 1.85) relative to those with adequate levels of vitamin D. Respondents with a serum vitamin D concentration less than 12 ng/mL were also more likely to report cardiovascular disease (OR 1.87, 95% CI 1.28 to 2.74) relative to those with adequate levels of vitamin D before multivariate adjustment. After multivariate adjustment, respondents with serum vitamin D concentration between 12 and 20 ng/mL were still at increased risk of cardiovascular disease (OR 1.45, 95% CI 1.08 to 1.97), as were those with levels less than 12 ng/mL (OR 2.11, 95% CI 1.38 to 3.24). The multivariate logistic regression model included adjustment for age, gender, BMI, smoking, and physical activity. The model included interaction terms for age and gender, age

Table 14 Unadjusted and adjusted odds ratios for the association between serum vitamin D concentration and self-reported cardiovascular disease

Cardiovascular Disease		Odds Ratio	95% Wald Confidence Limits		P_{Trend}[1]
			Lower	Upper	
Unadjusted	Vitamin D >20 ng/mL	(Reference)	-	-	0.001
	Vitamin D 12-20 ng/mL	1.445	1.127	1.852	
	Vitamin D <12 ng/mL	1.874	1.282	2.741	
Adjusted[2]	Vitamin D >20 ng/mL	(Reference)	-	-	0.0011
	Vitamin D 12-20 ng/mL	1.453	1.075	1.965	
	Vitamin D <12 ng/mL	2.112	1.377	3.24	

[1] Significance of Wald chi square statistic
[2] Adjusted for age, gender, BMI, smoking, and physical activity. The model included interaction terms for age and gender, age and BMI, and age and smoking.

and BMI, and age and smoking. There was also evidence of a significant trend in the severity of vitamin D deficiency in both the unadjusted and adjusted models ($p<0.0001$ and $p=0.0011$, respectively).

The unadjusted and adjusted odds ratios for the association between self-reported diabetes and serum vitamin D concentration are shown in table 15. Before adjustment for confounding variables, respondents with a serum vitamin D concentration between 12 and 20 ng/mL were significantly more likely to report diabetes (OR 1.76, 95% CI 1.41 to 2.21) relative to those with adequate levels of vitamin D. Respondents with a serum vitamin D concentration less than 12 ng/mL were more than twice as likely to report diabetes (OR 2.05, 95% CI 1.53 to 2.75) relative to those with adequate levels of vitamin D before multivariate adjustment. However, after multivariate adjustment there was no longer a significant association between vitamin D deficiency and self-reported diabetes. The multivariate logistic regression model included adjustment for age, race/ethnicity, poverty income ratio, BMI, video screen viewing and physical activity. There was evidence of a significant trend in the severity of vitamin D deficiency in the unadjusted model ($p<0.0001$) but not in the multivariate adjusted model.

Table 15 Unadjusted and adjusted odds ratios for the association between serum vitamin D concentration and self-reported diabetes

Diabetes		Odds Ratio	95% Wald Confidence Limits		P_{Trend}[1]
			Lower	Upper	
Unadjusted	Vitamin D >20 ng/mL	(Reference)	-	-	<.0001
	Vitamin D 12-20 ng/mL	1.764	1.406	2.213	
	Vitamin D <12 ng/mL	2.054	1.533	2.75	
Adjusted[2]	Vitamin D >20 ng/mL	(Reference)	-	-	0.2743
	Vitamin D 12-20 ng/mL	1.212	0.954	1.539	
	Vitamin D <12 ng/mL	1.089	0.804	1.475	

[1] Significance of Wald chi square statistic
[2] Adjusted for age, race/ethnicity, poverty income ratio, BMI, video screen viewing and physical activity.

Table 16 shows the unadjusted and adjusted odds ratios for the association between self-reported kidney disease and serum vitamin D concentration. Before adjustment for confounding variables, respondents with a serum vitamin D concentration between 12 and 20 ng/mL were almost twice as likely to report kidney disease (OR 1.98, 95% CI 1.22 to 3.21) relative to those with adequate levels of vitamin D. Respondents with a serum vitamin D concentration less than 12 ng/mL were more than twice as likely to report kidney disease (OR 2.15, 95% CI 1.20 to 3.84) relative to those with adequate levels of vitamin D before multivariate adjustment. There was a significant trend across vitamin D levels (p=0.0053). The multivariate logistic regression model included adjustment for age, poverty income ratio, season of survey administration and interaction between age and poverty income ratio. After multivariate adjustment, respondents with serum vitamin D concentration between 12 and 20 ng/mL were still at increased risk of kidney disease (OR 1.77, 95% CI 1.08 to 2.91), while those with levels less than 12 ng/mL were not (OR 1.74, 95% CI 0.92 to 3.32). The p value for trend was slightly outside the threshold for significance ($p < 0.05$) at p=0.0659.

Table 16 Unadjusted and adjusted odds ratios for the association between serum vitamin D concentration and self-reported kidney disease

Kidney Disease		Odds Ratio	95% Wald Confidence Limits		P_{Trend}[1]
			Lower	Upper	
Unadjusted	Vitamin D >20 ng/mL	(Reference)	-	-	0.0053
	Vitamin D 12-20 ng/mL	1.98	1.22	3.21	
	Vitamin D <12 ng/mL	2.15	1.20	3.84	
Adjusted[2]	Vitamin D >20 ng/mL	(Reference)	-	-	0.0659
	Vitamin D 12-20 ng/mL	1.77	1.08	2.91	
	Vitamin D <12 ng/mL	1.74	0.92	3.32	

[1] Significance of Wald chi square statistic
[2] Adjusted for age, poverty income ratio, season of survey administration and interaction between age and poverty income ratio.

The unadjusted and adjusted odds ratios for the association between self-reported osteo-porosis and serum vitamin D concentration are shown in table 17. There was no evidence of a significant association between serum vitamin D concentration and self-reported osteoporosis in either the unadjusted or multivariate adjusted models. The multivariate logistic regression model included adjustment for age, gender, race/ethnicity, BMI, video screen viewing, smoking and physical activity.

Table 17 Unadjusted and adjusted odds ratios for the association between serum vitamin D concentration and self-reported osteoporosis

Osteoporosis		Odds Ratio	95% Wald Confidence Limits		P_{Trend}[1]
			Lower	Upper	
Unadjusted	Vitamin D >20 ng/mL	(Reference)	-	-	0.631
	Vitamin D 12-20 ng/mL	0.9	0.653	1.241	
	Vitamin D <12 ng/mL	0.844	0.555	1.282	
Adjusted[2]	Vitamin D >20 ng/mL	(Reference)	-	-	0.9181
	Vitamin D 12-20 ng/mL	0.971	0.672	1.402	
	Vitamin D <12 ng/mL	0.893	0.521	1.53	

[1] Significance of Wald chi square statistic
[2] Adjusted for age, gender, race/ethnicity, BMI, video screen viewing, smoking and physical activity

4 DISCUSSION

4.1 Demographic and Behavioral Variables Associated with Serum 25(OH)-Vitamin D Concentration

Our findings related to the determinants of vitamin D status are for the most part consistent with those reported elsewhere in the literature (Forrest and Stuhldreher, 2011; Ganji, Zhang and Tanpricha, 2012; Ginde, Liu and Camargo, 2009; Melamed et al., 2009). In both the descriptive and multivariate analysis we found that race/ethnicity other than non-Hispanic White was associated with lower serum vitamin D levels. Non-Hispanic Black race/ethnicity, in particular, was strongly associated with lower levels of serum vitamin D in the multivariate analysis. These results are consistent with results reported elsewhere for the United States adult population (Forrest and Stuhldreher, 2011; Ganji, Zhang and Tanpricha, 2012; Ginde, Liu and Camargo, 2009; Melamed et al., 2009). Like Forrest and Stuhldreher (2011), we found that age was not significantly associated with vitamin D staus in the descriptive analysis. However, similar to Melamed, et al. (2009), we found that after multivariate adjustment age was a significant predictor for vitamin D concentration, and that younger age was associated with higher serum vitamin D levels in adults. As in the model-adjusted findings presented in Melamed, et al. (2009), we found that female gender, overweight and obesity (measured by BMI), and fall/winter seasonality were all associated with lower serum vitamin D, while supplement use and physical activity were associated with higher levels of vitamin D in our analysis. Although our study focused on adults, our finding of an inverse relationship between vitamin D and video screen viewing is consistent with the model adjusted results reported by Kumar et al., (2009) for children and adolescents between 1 and 21 years old. We found that neither poverty income group nor cigarette smoking were significantly associated with serum vitamin D concentration after multivariate adjustment, although these variables were significant in the unadjusted analysis.

4.2 Dietary Patterns and Vitamin D

The results of this study suggest that, after adjusting for potential confounders, healthy diets containing more servings of fruits, vegetables and lean meats are associated with higher mean serum vitamin D concentrations, while those with more servings of high fat meats and processed foods are associated with lower serum vitamin D concentrations. The regression analysis found a significant positive association between Dietary Pattern 1, characterized by fruits, vegetables and lean meats, and serum vitamin D concentration. Dietary Pattern 1 contained many foods that would be expected to be positively associated with serum vitamin D concentration, and this dietary pattern was significantly associated with vitamin D status after controlling for demographic, behavioral and environmental factors. In the principal components analysis, factor loadings for low fat dairy, fish and eggs were important but not dominant (i.e. factor loading > 0.5) contributors to total factor score for Dietary Pattern 1. Milk, which is typically fortified with vitamin D in the United States, and naturally occurring vitamin D in eggs and fish, are important dietary sources of vitamin D. However, the most dominant contributors to Factor 1 included fruits and vegetables, which were not associated with vitamin D status in our analysis or in the nutrition literature, and higher scores for this factor are driven most by more frequent consumption of fruits and vegetables rather than foods containing vitamin D. It is possible that this dietary pattern is associated with an overall healthier lifestyle, which may also lead to higher vitamin D levels through increased sun exposure as a result of outdoor physical activity.

We found a significant negative association between Dietary Pattern 2, characterized by high fat meats and processed foods, and serum 25(OH)-Vitamin D concentration after multivariate adjustment. It is important to note that the factor loading for low fat dairy was negative, which means that consumption of low fat milk would lead to slightly lower scores for this pattern. However, while there was a positive factor loading for high fat dairy, these foods may include vitamin D fortified high fat dairy products such whole milk, but also unfortified foods such as cheese. Overall, this dietary pattern was associated with foods that are generally considered unhealthy, and may be associated with other unhealthy lifestyle factors that were not captured by the variables in this study. It is not clear whether the association with lower vitamin D levels is due to the exclusion of healthy sources of vitamin D in favor of unhealthy foods, or whether the diet described by this pattern is characteristic of a more generally unhealthy lifestyle that would lead to less sun exposure and thus less endogenous vitamin D production.

The dietary patterns that we derived from the NHANES food frequency data using principal components analysis were similar to those found by other researchers using similar methods but different data sets. Ganji, Kafai and McCarthy (2009) also derived three components or dietary patterns with similar characteristic foods, including one pattern measuring intake of fruits vegetables and lean meat, another pattern measuring intake of red and processed meats and processed foods, and another pattern measuring a mixed dietary pattern. While our food categorization was similar to Ganji, Kafai and McCarthy (2009), we added additional categories to include condiments, dairy alternative meal replacement, other fats, other soups and other vegetables. The factor loadings for our dietary patterns for 1) fruits, vegetables and lean meat, and 2) high fat meats and processed foods were similar to those found by Ganji, Kafai and McCarthy (2009), but our third pattern described intake of condiments, butter or margarine and dairy alternative meal replacement rather than a mixed pattern of food consumption. Their three dietary patterns capture approximate 27.5% of the variation in their dietary data from NHANES III, which is somewhat less than our analysis which captured about one third of the dietary variation. Cho, Shin and Kim (2011) used principal components analysis to find dietary patterns using food groups appropriate for a Korean diet. Despite the difference in the underlying food groups, they also found three major dietary patterns that measured approximately 32% of the variation in their diet data.

4.3 Vitamin D and Self-Reported Medical Conditions

Before adjustment for potential confounders, we found that four self-reported medical conditions were associated with vitamin D status: cancer, cardiovascular disease, diabetes, and kidney disease. After multivariate adjustment we found that vitamin D deficiency was significantly related to increased risk for cardiovascular disease and kidney disease. We fitted a logistic regression model for each medical condition and included the dietary patterns as potential covariates along with the other study variables. However, contrary to our expectations the dietary factors were not significant predictors in any of the models for the five medical conditions.

The unadjusted results for self-reported cancer suggested that the likelihood of cancer was actually positively associated with serum vitamin D levels. In other words, respondents with lower serum vitamin D concentrations were less likely to report having cancer. It is not clear why vitamin D deficiency would be protective against cancer, although this relationship was not significant after adjusting for confounding variables. Since we considered all types of cancer, it is possible that this finding may have been influenced by the relatively high proportion of cancer cases (36.6%) represented by skin cancer. Overall, our results are consistent with those of other researchers in that we were unable to find substantial evidence of a link between cancer and vitamin D (Mocellin, 2011).

We found strong evidence for an associations between self-reported cardiovascular disease and vitamin D insufficiency and deficiency (serum 25(OH)-vitamin D concentration ≤ 20 ng/mL) even after multivariate adjustment. The trend in risk of cardiovascular disease significantly increased as serum vitamin D concentration decreased (p=0.0011). Many researches have noted a similar relationship between vitamin D and cardiovascular disease, and our results are consistent with those reported elsewhere (Leu and Giovannucci, 2011; Sun et al., 2011).

We also found evidence for a relationship between vitamin D and self-reported kidney disease even after adjusting for potential confounders. This finding makes sense after considering the role of the kidneys in vitamin D metabolism (Lehman and Meurer, 2010) and the previously discussed findings linking albuminuria to vitamin D deficiency (Fiscella, Winters and Ogedegbe, 2011). Mehrota et al., (2009) found an increased risk of mortality in persons with chronic kidney disease associated with serum vitamin D concentrations below 15 ng/mL, which roughly coincides with our cut point for increased risk of kidney disease (12-20 ng/mL).

Although other researchers have shown a link between type II diabetes risk and vitamin D deficiency in adults and children, our analysis found that the association between self-reported diabetes and vitamin D was not significant after controlling for the other study variables. This result is unexpected given the shared risk factors for both diabetes and cardiovascular disease. Several of these common risk factors often occur together and are known as cardiometabolic syndrome. Ganji, Shaikh, Zhang and Tangpricha (2011) found an association between low serum vitamin D levels and increased risk of cardiometabolic syndrome in U.S. adolescents aged 12 to 19 years old after controlling for potential confounders. They found that children in the lowest tertile for serum vitamin D were 1.71 times more likely (95% CI 1.11 to 2.65) to have metabolic syndrome than children in the highest category. Cardiometabolic syndrome encompasses several risk factors for cardiovascular disease and diabetes. Other authors have noted a similar inverse relationship between serum vitamin D concentration and diabetes risk among adults (Brock et al., 2011; Mitri, Muraru and Pittas, 2011).

The self-reported nature of the survey questions in the medical conditions part of the NHANES questionnaire almost certainly affected our findings. Survey participants were asked to respond to a question about whether a medical professional had ever told them that they had the medical condition, and this implies that the condition had already been clinically diagnosed and that the patient may have been receiving treatment. In effect, we were measuring the association between vitamin D status and whether a person had reported being clinically diagnosed with a medical condition. This type of response is subject to recall bias (Coughlin, 1990). Even if recall bias were not a factor, assuming a person had been diagnosed and was currently receiving treatment for the condition, the treatment could include drugs that affect vitamin D metabolism or may even include increased vitamin D intake. For example, we did not find an association between vitamin D status and self-reported osteoporosis, despite the well-known relationship between vitamin D deficiency and low bone mineral density (Bahn, Rao and Rao, 2010; Mauricic 2008). Since increased milk consumption or vitamin D supplementation is recommended for persons diagnosed with osteoporosis, persons who reported having had osteoporosis may have increased their dietary intake of vitamin D based on advice from their physician (Mauricic, 2008). Holick, et al. (2005) report a prevalence of 18% for vitamin D insufficiency or deficiency (serum 25(OH)-vitamin D less than 20 ng/mL) among a cohort of postmenopausal women who were receiving osteoporosis therapy. This proportion is still lower than the prevalence of 38.3% that we found for vitamin D insufficiency and deficiency (serum 25(OH)-vitamin D $\leq$ 20 ng/mL) in the general population.

4.4 Study Limitations

Our study is subject to some limitations. First, as with all health surveys our data set may have been subject to recall bias. Another limitation is that, although we were able to provide strong evidence of an association between vitamin D status and several chronic diseases, we are unable to establish a causal link between vitamin D deficiency and those medical conditions due to the cross-sectional design of the NHANES data set. NHANES also does not directly measure sun exposure, which in the primary source of vitamin D production for most people. Therefore we were unable to measure or control for the primary determinant of our outcome variable. This is especially important when studying dietary sources of vitamin D, since it is possible that sun exposure could have a confounding effect on dietary variables that are associated with outdoor activities or an active lifestyle.

Skin complexion is another important factor in endogenous vitamin D production is skin complexion (Hochberg and Templeton, 2010). Race/ethnicity may be a poor proxy for skin complexion since many biological and social scientists consider both race and ethnicity to be socially mediated categories of ambiguous biological significance (American Association of Physical Anthropologists, 1996). Researchers have noted that skin complexions may vary widely across persons of the same race/ethnicity group (American Association of Physical Anthropologists, 1996). While darker skin complexion may partially explain the association between race/ethnicity and vitamin D deficiency, there are other factors that should be considered. A high prevalence of lactose intolerance (lactase non-persistence) among race/ethnicity groups with historically non-pastoral food production patterns may discourage some people from consuming vitamin D fortified milk (Tishkoff et al., 2006). Also, although our dietary analysis captured approximately 35% of variation in diet, there remained 65% for which we still did not account. An analysis of dietary patterns that focuses specifically on foods rich in vitamin D may more effectively capture relevant dietary behaviors. Finally, although SAS 9.2 offers a full suite of statistical procedure for complex sample surveys, our analysis may be enhanced by using the advanced features available in SUDAAN version 10 software.

5 CONCLUSION

Vitamin D and its role in health and disease has been the subject of much recent research. Our analysis of nationally representative nutritional survey data from the National Health and Nutrition Examination Survey for the years 2003-2006 has examined the importance of several key demographic, behavioral, environmental and nutritional variables in maintaining healthy levels of vitamin D in the body. We found that serum 25(OH)-Vitamin D concentration was positively associated with diets high in fruits, vegetables, lean meats and low-fat dairy, while diets high in processed foods and high-fat meats were inversely associated with vitamin D level. We also found that older age, non-White race/ethnicity, higher than normal BMI and fall/winter seasonality were inversely associated with serum vitamin D concentration. Supplementation, however, contributed significantly to higher serum levels of vitamin D. Finally, after controlling for potential confounders, vitamin D insufficiency or deficiency was positively associated with two self-reported medical conditions: cardiovascular disease and kidney disease.

6 REFERENCES

American Association of Physical Anthropologists. (1996). "AAPA Statement on Biological Aspects of Race." *American Journal of Physical Anthropology* 101:569-570.

Brock KE, Huang WY, Fraser DR, Kea L, Tseng M, Mason RS, Stolzenberg-Solomon RZ, et al. (2011.) "Diabetes prevalence is associated with serum 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D in US middle-aged Caucasian men and women: a cross-sectional analysis within the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial." *British Journal of Nutrition* 106:339-344.

Centers for Disease Control and Prevention (CDC). (2003). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2003 http://www.cdc.gov/nchs/nhanes/nhanes2003-2004/nhanes03_04.htm.

Centers for Disease Control and Prevention (CDC). (2005). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2005 http://www.cdc.gov/nchs/nhanes/nhanes2005-2006/nhanes05_06.htm.

Centers for Disease Control and Prevention (CDC). (2006). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Analytical Guidelines. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, http://www.cdc.gov/nchs/nhanes/nhanes2003-2004/analytical_guidelines.htm

Cho YA, Chin A, Kim J. (2011). "Dietary patterns are associated with body mass index in a Korean population." *Journal of the American Dietetic Association* 111:1182-1186.

Coughlin SS. (1990). "Recall bias in epidemiologic studies." *Journal of Clinical Epidemiology* 43(1):87-91

Fiscella KA, Winters PC, and Ogedegbe G. (2011). Vitamin D and racial disparity in albuminuria: NHANES 2001-2006. *American Journal of Hypertension* 24(10):1114-1120.

Forrest KY, Stuhldreher WL. (2011). "Prevalence and correlates of vitamin D deficiency in US adults." *Nutr Res* Jan, 31(1):48-54.

Freedman DM, Looker AC, Abnet CC, Linet SM, and Graubard BI. (2010). "Serum 25-hydroxyvitamin D and cancer mortality in the NHANES III study (1988-2006)." *Cancer Res* 70(21):8587-97.

Gallieni M, Cozzolino M, Fallabrino G, Pasho S, Olivi L, Brancaccio D. (2009). "Vitamin D: physiology and pathophysiology." *Int J Artif Organs* 32:87-94.

Ganji V, Milone C, Cody MM, McCarty F, Wang YT. (2010). "Serum vitamin D concentrations are related to depression in young adult US population: the Third National Health and Nutrition Examination Survey. " *Int Arch Med* 3:29.

Ganji V, Kafai MR, and McCarthy E. (2009). "Serum leptin concentrations are not related to dietary patterns but are related to sex, age, body mass index, serum triacylglycerol, serum insulin, and plasma glucose in the US population." *Nutrition and Metabolism* 6:3.

Ganji V, Zhang X, Shaikh N, Tangpricha V. (2011). "Serum 25-hydroxyvitamin D concentrations are associated with prevalence of metabolic syndrome and various cardiometabolic risk factors in US children and adolescents based on assay-adjusted serum 25-hydroxyvitamin D data from NHANES 2001-2006." *Am J Clin Nutr* 94:225-233.

Ganji V, Zhang X, Tangpricha V. (2012). "Serum 25-hydroxyvitamin D concentrations and prevalence estimates of hypovitaminosis D in the U.S. population based on assay-adjusted data." *J Nutr* 142(3):498-507.

Ginde AA, Liu MC, and Camargo CA. (2009). "Demographic differences and trends of vitamin D insufficiency in the US population, 1988-2004. " *Arch Intern Med* Mar 23, 169(6):626-32.

Hochberg Z and Templeton AR. (2010). "Evolutionary perspective in skin color, vitamin D and its receptor." *Hormones* 9(4):307-311.
Holick, MF, Siris ES, Binkley N, Beard MK, Khan A, Katzer JT, et al. (2005). "Prevalence of vitamin D inadequacy among postmenopausal North American women receiving osteoporosis therapy." *Journal of Clinical Endocrinology and Metabolism*. 2005; 90(6): 3215-3224.

Institute of Medicine. Food and Nutrition Board. (2010). *Dietary Reference Intakes for Calcium and Vitamin D.* Washington, DC: National Academy Press.

Kumar J, Muntner P, Kaskel FJ, Hailpern SM, Melamed ML. (2009). "Prevalence and associations of 25-hydroxyvitamin D deficiency in US children: NHANES 2001-2004." *Pediatrics* 124:e362-70.

Larsen ER, Mosckilde L and Foldspang A. (2004). "Vitamin D and calcium supplementation prevents osteoporotic fractures in elderly community dwelling residents: a pragmatic population-based 3-year intervention study." *Journal of Bone and Mineral Research* 19: 370-378.

Lehmann B, Meurer M. (2010). "Vitamin D metabolism." *Dermatol Ther* 23:2-12.

Leu, M, and Giovannucci E. (2011). "Vitamin D: Epidemiology of cardiovascular risks and events. Best Practice and Research Clinical Endocrinology and Metabolism." 25: 633-646.

Llewellyn, DJ, Lang IA, Langa KM, and Melzer D. (2011). "Vitamin D and cognitive impairment in the elderly U.S. population." *J Gerontol Biol Sci Med Sci* January; 66A(1): 59-65.

Lohr, S. (1999). *Sampling Design and Analysis*. Boston: Duxbury Press.

Ma Y, Zhang P, Wang F, Yang J, Liu Z, and Qin H. (2011). "Association between vitamin D and risk of colorectal cancer: A systematic review of prospective studies." *Journal of Clinical Oncology*. Available at http://jco.ascopubs.org/content/early/2011/08/28/JCO.2011.35.7566.full.pdf+html.

Makariou S, Liberopoulos EN, Elisaf M, and Challa A. (2011). "Novel roles of vitamin D in disease: What is new in 2011?" *European Journal of Internal Medicine* 22: 355-362.

Mansbach JM, Ginde AA, and Camargo CA. (2009). "Serum 25-hydroxyvitamin D levels among US children aged 1 to 11 years: do children need more vitamin D?" *Pediatrics* Nov;124(5):1404-10.

Maricic M. (2008). "Osteomalacia." *Current Osteoporosis Reports*, 6: 130-133.

Melamed, ML, B Astor, ED Michos, TH Hostetter, NR Powe, and P Muntner. (2009). " 25-hydroxyvitamin D levels, race, and the progression of kidney disease." *J Am Soc Nephrol* 20: 2631-2639.

Mitri, J, MD Muraru and AG Pittas. (2011). "Vitamin D and type 2 diabetes: a systematic review." *European Journal of Clinical Nutrition* 65: 1005-1015.

Rao, JNK., and Scott AJ.(1979), "Chi-Squared Tests for Analysis of Categorical Data from Complex Surveys, " *Proceedings of the Survey Research Methods Section*, ASA, 58–66.

SAS Institute, Inc. (2012). Estimating Geometric Means Using Data from a Complex Survey Sampling Design. Available at: http://support.sas.com/rnd/app/examples/stat/SurveyGeoMean/stat_webex.pdf. Accessed May 4, 2012.

Mocellin S. (2011) "Vitamin D and cancer: deciphering the truth." *Biochimica et Biophysica Acta* 172-178.

Montgomery DC, Peck AP and Vining GG. (2006). *Introduction to Linear Regression Analysis.* Hoboken, NJ: Wiley.

Norman AW. (2008). "From vitamin D to hormone D: fundamentals of the vitamin D endocrine system essential for good health." *Am J Cl Nutr* 88:491S-99S.

Rajakumar K, Greenspan SL, Thomas SB, and Holick MF. (2007). "Solar ultraviolet radiation and vitamin D." *Am J Public Health* 97: 1740-1748.

Research Triangle Institute. (2005). Sudaan version 10 software for statistical analysis. Research Triangle Park, NC: RTI.

Sun Q, Shi L, Rimm EB, Giovannucci EL, Hu FB, et al. (2011). “Vitamin D intake and risk of cardiovascular disease in US men and women.” *Am J Clin Nutr* 94;534-42.

Tishkoff, SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, et al. (2007). “Convergent adaptation of human lactase persistence in Africa and Europe.” *Nature Genetics* 39: 31-40.

U.S. Department of Agriculture, Agricultural Research Service. 2011. “USDA National Nutrient Database for Standard Reference, Release 24.” Nutrient Data Laboratory Home Page,http://www.ars.usda.gov/ba/bhnrc/ndl

Zittermann A, Schleithoff SS and Koerfer R. (2005). “Putting cardiovascular disease and vitamin D insufficiency into perspective.” *British Journal of Nutrition* 94, 483–492.

7 APPENDIX

FFQ FOOD	VALUE	FOOD GROUP	DESCRIPTION
1	Milk whole in cereal	2	High fat dairy
2	Milk 2 pct in cereal	1	Low fat dairy
3	Milk 1 pct in cereal	1	Low fat dairy
4	Milk nonfatskim in cereal	1	Low fat dairy
5	Milk soy in cereal	3	Dairy AltMeal Repl
6	Milk rice in cereal	3	Dairy AltMeal Repl
7	Non-dairy crm powdrd reg in coffee or tea	3	Dairy AltMeal Repl
8	Non-dairy crm powdrd diet in coffee or tea	3	Dairy AltMeal Repl
9	Non-dairy crm liquid reg in coffee or tea	3	Dairy AltMeal Repl
10	Non-dairy crm liquid diet in coffee or tea	3	Dairy AltMeal Repl
11	Cream reg or 1212 in coffee or tea	2	High fat dairy
12	Milk whole to drink	2	High fat dairy
13	Milk 2 pct to drink	1	Low fat dairy
14	Milk 1 pct to drink	1	Low fat dairy
15	Milk nonfat to drink	1	Low fat dairy
16	Milk soy to drink	3	Dairy AltMeal Repl
17	Milk rice to drink	3	Dairy AltMeal Repl
18	Milk whole in coffee or tea	2	High fat dairy
19	Milk 2 pct in coffee or tea	1	Low fat dairy
20	Milk 1 pct in coffee or tea	1	Low fat dairy
21	Milk nonfatskim in coffee or tea	1	Low fat dairy
22	Milk evapcond in coffee or tea	1	Low fat dairy
23	Milk soy in coffee or tea	3	Dairy AltMeal Repl
24	Milk rice in coffee or tea	3	Dairy AltMeal Repl
25	Meal repl liquid	3	Dairy AltMeal Repl
26	Yogurt all	1	Low fat dairy
27	Cotricot cheese	2	High fat dairy
28	Sour cream reg	2	High fat dairy
29	Sour cream lowfat	1	Low fat dairy
30	Cheese reg	2	High fat dairy
31	Cheese lowfat	1	Low fat dairy
32	Cream cheese reg	2	High fat dairy
33	Cream cheese lowfat	1	Low fat dairy
34	Ice cream reg	2	High fat dairy
35	Ice creamice milk lowfat	1	Low fat dairy
36	Frozen yogurt ices sorbet etc	1	Low fat dairy
37	Puddingscustards	2	High fat dairy

FFQ FOOD	VALUE	FOOD GROUP	DESCRIPTION
38	Beef steaks reg	6	Meat
39	Beef steaks lean	6	Meat
40	Beef roast	6	Meat
41	Roast beef in sandwich	6	Meat
42	Beef stewspot piesmixtures	12	Mixed dishes
43	Beef burgers lean	6	Meat
44	Beef burgers reg	6	Meat
45	Beef gr meatballsloafmixtures	6	Meat
46	Ham not luncheon	6	Meat
47	Pork	6	Meat
48	Bacon regular	7	Processed Meat
49	Bacon leanCanadian	7	Processed Meat
50	Sausage reg	7	Processed Meat
51	Sausage turklowfat	7	Processed Meat
52	Hot dogs regular	7	Processed Meat
53	Hot dogs turkylowfat	7	Processed Meat
54	Shortribsspareribs	6	Meat
55	Liver liverwurst	7	Processed Meat
56	Cold cuts regular	7	Processed Meat
57	Cold cuts lowfat	7	Processed Meat
58	Cold cuts poultry	7	Processed Meat
59	Ham cold cut lunch meat reg	7	Processed Meat
60	Ham cold cut lunch meat lowfat	7	Processed Meat
61	Chicken fr light wskin	8	Poultry
62	Chicken fr light woskin	8	Poultry
63	Chicken fr dark wskin	8	Poultry
64	Chicken fr dark woskin	8	Poultry
65	Chicken light wskin	8	Poultry
66	Chicken light woskin	8	Poultry
67	Chicken dark wskin	8	Poultry
68	Chicken dark woskin	8	Poultry
69	Chickenturkey ground	8	Poultry
70	Chicken mixtures	12	Mixed dishes
71	Turkey	8	Poultry
72	Tofu soy meats	17	Legumes
73	Tuna canned	4	Fish other seafood
74	Fish fried	4	Fish other seafood
75	Fish not fried	4	Fish other seafood
76	Fish oysters	4	Fish other seafood
77	Eggs regular	5	Eggs

FFQ FOOD	VALUE	FOOD GROUP	DESCRIPTION
78	Eggs whites only	5	Eggs
79	Eggs substitutes	5	Eggs
80	Eggs salad	5	Eggs
81	Soups broth w ndlesrice	10	Other soup[1]
82	Soups w veggies	10	Other soup[1]
83	Soups bean-type	10	Other soup[1]
84	Soups creamed	9	Creamed soup[1]
85	Eng mufbagel	14	Refined grains
86	Breadsrolls white	14	Refined grains
87	Breadnot white	15	Whole grains
88	Crackers	14	Refined grains
89	Stuffingdumplings	14	Refined grains
90	Cornbreadmuffins	14	Refined grains
91	Biscuits all	14	Refined grains
92	Donuts swt rolls danish pop tarts	24	Sweets Snacks
93	Muffinsdessert breads	24	Sweets Snacks
94	Cookies brownies	24	Sweets Snacks
95	Granola bars	15	Whole grains
96	Cakes	24	Sweets Snacks
97	Pies fruit	24	Sweets Snacks
98	Pies, crmcustrdothr	24	Sweets Snacks
99	Crispscobblers	24	Sweets Snacks
100	Popcorn	24	Sweets Snacks
101	Pretzels	24	Sweets Snacks
102	Nutsseeds whole	16	Nuts
103	Nutsseeds butters	16	Nuts
104	Pancke waff Fr tst	14	Refined grains
105	RTE cereal<half whole grain	13	Cereal
106	Ricegrains white	14	Refined grains
107	Pasta no fat added	14	Refined grains
108	Pasta fat added	14	Refined grains
109	Pasta meatless red sauce	12	Mixed dishes
110	Pasta meatfish sauce	12	Mixed dishes
111	Lasagna rav shells etc	12	Mixed dishes
112	Macaroni and cheese	12	Mixed dishes
113	Pasta salad	12	Mixed dishes
114	Pizza with meat	11	Pizza
115	Pizza without meat	11	Pizza
116	Oranges tangelo etc	22	Fruit
117	Grapefruit	22	Fruit

FFQ FOOD	VALUE	FOOD GROUP	DESCRIPTION
118	Apples	22	Fruit
119	Applesauceckd apples	22	Fruit
120	Pears	22	Fruit
121	Peachesnectarinesplums	22	Fruit
122	Bananas	22	Fruit
123	Melons	22	Fruit
124	Strawberries	22	Fruit
125	Grapes all	22	Fruit
126	Dried fruit	22	Fruit
127	Other fruits	22	Fruit
128	Orangegrpfrt jce all	23	Fruit juices
129	Other juice	23	Fruit juices
130	Tomatoveg juice all	19	Tomatoes
131	Beans	17	Legumes
132	Chili	12	Mixed dishes
133	Potatoes white, no fat added	18	Starchy vegetables
134	Potatoes fried	18	Starchy vegetables
135	Potato salads	18	Starchy vegetables
136	Sweet potatoes, no fat added	18	Starchy vegetables
137	Pickled vegfrt	21	Other vegetables
138	Raw spinachgreens	20	Cruciferous green vegetables
139	Ckd spinachgreens, no fat added	20	Cruciferous green vegetables
140	Broccoli, no fat added	20	Cruciferous green vegetables
141	Carrots, no fat added	18	Starchy vegetables
142	Tomatoes raw	19	Tomatoes
143	Tomato salsa	19	Tomatoes
144	Tomato catsup	19	Tomatoes
145	String beans, no fat added	21	Other vegetables
146	Cabbagesauerkraut	20	Cruciferous green vegetables
147	Coleslaw	20	Cruciferous green vegetables
148	Peas, no fat added	21	Other vegetables
149	Corn, no fat added	22	Fruit
150	CauliflBr Spr, no fat added	20	Cruciferous green vegetables
151	Peppers, no fat added	21	Other vegetables
152	Onions, no fat added	21	Other vegetables
153	Veg med, no fat added	21	Other vegetables
154	Other vegetables, no fat added	21	Other vegetables
155	Margarine reg on breadpanwaff	25	ButterMargerine
156	Margarine low-fat on breadpanwaff	25	ButterMargerine
157	Butter reg on breadpanwaff	25	ButterMargerine

FFQ FOOD	VALUE	FOOD GROUP	DESCRIPTION
158	Butter reduced fat on breadpanwaff	25	ButterMargerine
159	Margarine reg on potveggrains	25	ButterMargerine
160	Margarine diet on potveggrains	25	ButterMargerine
161	Butter reg on potveggrains	25	ButterMargerine
162	Butter reduced fat on potveggrains	25	ButterMargerine
163	Oils olive	26	Other fats
164	Oils corn	26	Other fats
165	Oils canola	26	Other fats
166	Oils other	26	Other fats
167	Salad drsg all on salad or veg	26	Other fats
168	Mayonnaise reg	26	Other fats
169	Mayonnaise diet	26	Other fats
170	Maple syrup on pancakes etc	27	Added Sugars
171	Sugarshoney not in coffeetea	27	Added Sugars
172	Sugarshoney all in coffee or tea	27	Added Sugars
173	Candy chocolate	24	Sweets Snacks
174	Candy not chocolate	24	Sweets Snacks
175	Jams jelly frt butters	27	Added Sugars
176	Coffee reg no crsug	28	CoffeeTea
177	Coffee decaf no crsug	28	CoffeeTea
178	Tea reg no crsug	28	CoffeeTea
179	Tea decaf no crsug	28	CoffeeTea
180	Soft drinks reg caff	29	High energy drinks
181	Soft drinks reg decaf	29	High energy drinks
182	Soft drinks diet caff	29	Low energy drinks
183	Soft drinks diet decaf	29	Low energy drinks
184	Fruit drinks reg	29	High energy drinks
185	Frt drinks diet	29	Low energy drinks
186	Alc bev liquor	30	Alcohol
187	Beer	30	Alcohol
188	Wine	30	Alcohol
189	Gravy	26	Other fats
190	Artificial sweetener in coffeetea	27	Added Sugars
191	Apple juice	23	Fruit juices
192	Grape juice	23	Fruit juices
193	Milk, unpasteurized not in coffeetea	2	High fat dairy
194	Milk, unpasteurized in cereal	2	High fat dairy
195	Milk, unpasteurized in coffeetea	2	High fat dairy
196	Oatmeal	13	Cereal
197	Hot brkfst cereals (not oatmeal)	13	Cereal

FFQ FOOD	VALUE	FOOD GROUP	DESCRIPTION
198	RTE cereal> half whole grain	13	Cereal
199	Pineapple	22	Fruit
200	Cucumbers	21	Other vegetables
201	Squash	18	Starchy vegetables
202	Lettuce dark green	20	Cruciferous green vegetables
203	Lettuce, not dark green	20	Cruciferous green vegetables
204	Tortillastacos corn	24	Sweets Snacks
205	Tortillastacos wheat	24	Sweets Snacks
206	Ricegrains whlgrn	15	Whole grains
207	Fish smoked	4	Fish other seafood
208	Sushi raw fish	4	Fish other seafood
209	Sushi no raw fish	12	Mixed dishes
210	Potatoothr chips (not corn) reg	18	Starchy vegetables
211	Potatoothr chips (not corn) lowfat	18	Starchy vegetables
212	Corn chips reg	24	Sweets Snacks
213	Corn chips lowfat	24	Sweets Snacks
214	Milk other to drink	3	Dairy AltMeal Repl
215	Milk other in cereal	3	Dairy AltMeal Repl
216	Milk other in coffeetea	3	Dairy AltMeal Repl