

**National Health and Nutrition
Examination Survey 2005–2006**

**Documentation, Codebook,
and Frequencies**

**Food Frequency Questionnaire –
Output from DietCalc Software**

Examination

**Survey Years:
2005 to 2006**

**SAS Transport File:
FFQDC_D.XPT**



September 2008

NHANES 2005–2006 Data Documentation

Questionnaire Section: Post-MEC Exam

Name: FFQDC_D

Years of Coverage: 2005–2006

First Published: September 2008

Last Revised: NA

Component Description

A food frequency questionnaire (FFQ) (sometimes referred to as the NHANES food propensity questionnaire or “FPQ”) was used in NHANES 2003-2004 and 2005-2006 to collect information on the frequency of food consumption during the past 12 months.

Two public data release files were prepared. The FFQ Daily Frequency File (SAS name FFQDC_D) described here provides daily food frequency covariates. The second file, the FFQ Questionnaire File (SAS name: FFQRAW_D) is described separately. The FFQ data augment the other NHANES 2005-2006 dietary assessment components which include two 24-hour dietary recall interviews and interview information on dietary supplement use, food security, and dietary behavior.

FFQ Development And Testing

The NHANES FFQ questionnaire was developed by the National Institutes of Health, National Cancer Institute (NCI). The basis for the NHANES FFQ is the NCI Diet History Questionnaire (DHQ), a 124-item food frequency instrument that is widely used in nutritional epidemiology research (Subar 2001a). Unlike the DHQ, portion size information was not collected with the NHANES FFQ. The NHANES FFQ data were not intended to be used to derive estimates of absolute intake for either nutrients or foods. Diet*Calc software was developed by the National Cancer Institute (Diet*Calc 2005). Diet*Calc was used to produce daily food frequency estimates from the FFQ data.

Eligible Sample

All English or Spanish-speaking examinees 2+ years of age who completed at least one 24-hr dietary recall interview were eligible for the FFQ component.

Interview Setting and Mode of Administration

Printed FFQ questionnaires were mailed to survey participants’ homes. Proxy respondents (usually a parent) completed the questionnaire for children less than 6 years of age and proxies assisted children 6-11 years of age and persons who could not complete the questionnaire on their own. Subjects 12 years of age and over self-reported. A postage-paid envelope was provided to respondents. Respondents who returned their FFQ form received \$30 remuneration. The FFQ forms were scanned at a central office and the scanned data were added to the

NHANES database.

The Diet*Calc software was used with the scanned FFQ data to produce daily frequencies for multiple foods and food groups. The FFQDC_D file contains more than 790,000 records. A majority of survey participants have more than 100 Diet*Calc records.

Quality Assurance & Quality Control

Quality control programs and manual verification checks were used to verify the completeness and accuracy of the scanned FFQ data files. All of the scanned FFQ records were included in the questionnaire file (FFQRAW_D). The questionnaire file documentation describes the methodology which was used to determine the completeness of the FFQ data.

Data Processing and Editing

None of the completed FFQ questionnaires or data were edited.

Analytic Notes

Data users should review the NHANES Analytic Guidelines carefully before analyzing the FFQ data. Additional guidance will be provided at a later date when NHANES website tutorials on NHANES dietary data analysis are available.

Two look-up tables are provided for use with the FFQ. The look-up tables (file names: FOODLOOK and VARLOOK) provide descriptive text labels for the food and variable identification number variables in this file.

A SAS program is provided on the NHANES website to show how to combine the descriptive text and frequency data.

The daily frequencies in this file are based on algorithms within Diet*Calc. These algorithms assign a daily frequency as follows:

For beverages:

Never = 0

1 time per month or less = 0.03

2-3 times per month = 0.08

1-2 times per week = 0.21

3-4 times per week = 0.5

5-6 times per week = 0.79

1 time per day = 1

2-3 times per day = 2.5

4-5 times per day = 4.5
6 or more times per day = 7

For foods:

Never= 0
1-6 time per year = 0.01
7-11 times per year = 0.028
1 time per month = 0.033
2-3 times per month = 0.08
1 time per week = 0.14
2 times per week = 0.29
3-4 times per week = 0.5
5-6 times per week = 0.79
1 time per day = 1
2 or more times per day = 2

For coffee, tea, and additions to coffee and tea:

None= 0
Less than 1 cup per month = 0.02
1-3 cups per month = 0.07
1 cup per week = 0.14
2-4 cups per week = 0.43
5-6 cups per week = 0.79
1 cup per day = 1
2-3 cups per day = 2.5
4-5 cups per day = 4.5
6 or more cups per day = 7

Proportion Question Algorithms: The proportion question(s) which accompany a particular frequency question will also affect the daily frequencies. The proportion (P) will act as a multiplier for the specific type of food mentioned in the proportion question, and the remaining proportion (1-P) is used as the multiplier for the food not specifically mentioned in the proportion question. For instance, the proportion question for question 6 would apply the answer supplied as a proportion of diet or sugar-free fruit drinks, while the remaining proportion would be applied to regular fruit drinks. The multipliers are obtained using the following algorithm:

Almost never or never = 0
About $\frac{1}{4}$ of the time = 0.25
About $\frac{1}{2}$ of the time = 0.5

About ¾ of the time = 0.75
 Almost always or always = 1

However, the proportion questions for questions 9, 84 use a different algorithm than the one above. For questions 9 and 84, each proportion can apply simultaneously for those foods. Using question 9 as an example, here is how the algorithm would work:

First proportion (how often soda was diet or sugar-free) = P1
 Second proportion (how often soda was caffeine-free) = P2

Multiplier used for:
 Diet, caffeine-free soda = P1 * P2
 Diet, caffeinated soda = P1 * (1 – P2)
 Regular, caffeine-free soda = (1 – P1) * P2
 Regular, caffeinated soda = (1 – P1) * (1 – P2)

Seasonal Question Algorithms: for seasonal questions (9, 10, 13, 21, 23, 24, 25, 26, 35, 42, 43, 97), the following algorithm was used to assign missing values:

| | If answer to lead-in question (i.e. Q9) is: | | |
|---------------------------------------|---|-----|---------|
| | No | Yes | Missing |
| Embedded questions (i.e. Q9a and Q9b) | Number of frequencies counted as missing | | |
| a and b missing | 0 | 2 | 2 |
| a and b not missing | 0 | 0 | 0 |
| a or b missing | 1 | 1 | 1 |

Several references describing the FFQ development and validation, the NCI statistical model, and appropriate uses of FFQ information are included in the References.

Special FFQ sample weights were produced for the FFQ subsample. Analysts should use the FFQ sample weights (WTS_FFQ) for all FFQ data analyses. The FFQ sample weights were derived from the NHANES 2005-2006 examined sample weights and were designed to account for FFQ non-response. Respondents with fewer than ten missing frequency values (i.e. FFQ_MISS < 10) have an FFQ sample weight.

The NHANES Analytic Guidelines describe the recommended procedures for analyzing the NHANES data properly.

Data Uses

NCI, in collaboration with USDA/CNPP, has developed models that can be used to estimate population usual intake distributions and examine relationships between dietary intake and health outcomes. For both of these uses, the models depend primarily on data from 24-hour recalls, but additional covariates may be employed to improve the estimates. These FFQ data were collected specifically to be used to help estimate usual intakes of episodically consumed foods. The models are based on the idea that usual food intake can be expressed as the probability of consuming the food on a given day multiplied by the usual amount consumed on consumption days. While 24-hour recall data provide detailed information about the amount consumed on consumption days, they provide relatively little information about the probability to consume. Therefore the FFQ-based average daily frequency values from file FFQDC_D may provide useful covariates regarding probability to consume and hence, lead to improved predictions of usual intake. Here, “improved predictions” means that more of the variation in usual intake can be explained by observed data. Partitioning the variance of usual intake into explained and unexplained components is a key step in evaluating relationships between usual dietary intake and health outcomes using the statistical method known as regression calibration. However, when estimating a population distribution of usual intake, it is not required to partition the variance of usual intake, but rather to estimate it in total. Therefore, using FFQ-based covariates does not necessarily improve estimation for population usual intake distributions, and therefore may not be required for that purpose.

Furthermore, it is not recommended that these FFQ data be used alone to estimate absolute intakes of foods or nutrients. Such data have been shown to exhibit significant measurement error, making them an inappropriate choice in this dataset that includes 24-hour recall data. For that reason, nutrient values are not provided.

References

CDC. NHANES Analytic and Reporting Guidelines. Updated September 2006. URL: http://www.cdc.gov/nchs/about/major/nhanes/nhanes2003-2004/analytical_guidelines.htm

CDC. NHANES 2005-2006 Website. URL: http://www.cdc.gov/nchs/about/major/nhanes/nhanes2005-2006/nhanes05_06.htm

Diet*Calc Analysis Program, Version 1.4.3. National Cancer Institute, Applied Research Program. November 1, 2005.

Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, Tooze JA, Krebs-Smith SM. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *J Am Diet Assoc.* 2006 Oct;106(10):1640-50. PMID: 17000197

NIH/NCI. Diet*Calc Analysis Program, Version 1.3.2. National Cancer Institute, Applied Research Program. June 2003. URL: <http://riskfactor.cancer.gov/DHQ/dietcalc/index.html>

Subar AF, Thompson FE, Smith AF, Jobe JB, Ziegler RG, Potischman N, Schatzkin A, Hartman A, Swanson C, Kruse L, Hayes RB, Lewis DR, Harlan LC. Improving food frequency questionnaires: a qualitative approach using cognitive interviewing. *J Am Diet Assoc* 1995;95:781-788.

Subar AF, Thompson FE, Kipnis V, Midthune D, Hurwitz P, McNutt S, McIntosh A, Rosenfeld S. Comparative Validation of the Block, Willett, and National Cancer Institute Food Frequency Questionnaires: The Eating at America's Table Study. *Am J Epidemiol* 2001a;154:1089-99.

Subar AF, Ziegler RG, Thompson FE, Johnson CC, Weissfeld JL, Reding D, Kavounis KH, Hayes RB for the Prostate, Lung, Colorectal, and Ovarian Cancer Screen Trial Investigators. Is shorter always better? Relative importance of questionnaire length and cognitive ease on response rates and data quality for two dietary questionnaires. *Am J Epidemiol* 2001b;153(4):404-9.

Subar AF, Dodd KW, Guenther PM, Kipnis V, Midthune D, McDowell M, Tooze JA, Freedman LS, Krebs-Smith SM. The food propensity questionnaire: concept, development, and validation for use as a covariate in a model to estimate usual food intake. *J Am Diet Assoc.* 2006 Oct;106(10):1556-63. PMID: 17000188

Thompson FE, Subar AF, Brown CC, Smith AF, Sharbaugh CO, Jobe JB, Mittl E, Gibson T, Ziegler RG. Cognitive research enhances accuracy of food frequency questionnaire reports: results of an experimental validation study. *J Am Diet Assoc* 2002 Feb;102(2):212-25.

Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, Guenther PM, Carroll RJ, Kipnis V. A new statistical method for estimating the usual intake of episodically consumed foods with

application to their distribution. J Am Diet Assoc. 2006 Oct;106(10):1575-87. PMID: 17000190

Locator Record

Title: NHANES Food Frequency Questionnaire Daily Frequency Covariates File

Contact Number: 1-866-441-NCHS

Years of Content: 2005–2006

First Published: September 2008

Revised: NA

Access Constraints: None

Use Constraints: None

Geographic Coverage: National

Subject: Food Frequency Questionnaire

Record Source: NHANES 2005–2006

Survey Methodology: NHANES 2005–2006 is a stratified multistage probability sample of the civilian non-institutionalized population of the U.S.

Medium: NHANES Web site; SAS transport files

**National Health and Nutrition Examination Survey
Codebook for Data Production (2005-2006)**

**Food Frequency Questionnaire Data - Output from DietCalc
(FFQDC_D)
Person Level Data**

September 2008



| | |
|--|----------------------------|
| SEQN | Target |
| | B(2 Yrs. to 150 Yrs.) |
| Hard Edits | SAS Label |
| | Respondent sequence number |
| English Text: Respondent sequence number. | |
| English Instructions: | |

| | |
|--|-------------------------------|
| FFQ_VAR | Target |
| | B(2 Yrs. to 150 Yrs.) |
| Hard Edits | SAS Label |
| | The Variable ID from DietCalc |
| English Text: The Variable ID from DietCalc. Use the VARLOOK dataset to cross-reference this Variable ID value to its corresponding food description. | |
| English Instructions: | |

| Code or Value | Description | Count | Cumulative | Skip to Item |
|---------------|-----------------|---------|------------|--------------|
| 1 to 152 | Range of Values | 1019121 | 1019121 | |
| . | Missing | 0 | 1019121 | |

| | |
|---|---------------------------|
| FFQ_FOOD | Target |
| | B(2 Yrs. to 150 Yrs.) |
| Hard Edits | SAS Label |
| | The Food ID from DietCalc |
| English Text: The Food ID from DietCalc. Use the FOODLOOK dataset to cross-reference this Food ID value to its corresponding food description. | |
| English Instructions: | |

| Code or Value | Description | Count | Cumulative | Skip to Item |
|---------------|-----------------|---------|------------|--------------|
| 1 to 216 | Range of Values | 1019121 | 1019121 | |
| . | Missing | 0 | 1019121 | |

| | | | | |
|------------------------------|--------------------------------------|--------------|-------------------|---------------------|
| FFQ_FREQ | Target | | | |
| | B(2 Yrs. to 150 Yrs.) | | | |
| Hard Edits | SAS Label | | | |
| | Daily intake frequency for this food | | | |
| English Text: | | | | |
| English Instructions: | | | | |
| Code or Value | Description | Count | Cumulative | Skip to Item |
| 0 to 7 | Range of Values | 1019121 | 1019121 | |
| . | Missing | 0 | 1019121 | |

| | | | | |
|--|---|--------------|-------------------|---------------------|
| FFQ_CODE | Target | | | |
| | B(2 Yrs. to 150 Yrs.) | | | |
| Hard Edits | SAS Label | | | |
| | Daily intake frequency imputation code | | | |
| English Text: Code describing how the food frequencies were derived. A value of zero means the daily food frequency was computed from the actual questionnaire response while values of 1, 2, and 3 represent imputed daily food frequencies. | | | | |
| English Instructions: | | | | |
| Code or Value | Description | Count | Cumulative | Skip to Item |
| 0 | Daily food frequency based on actual questionnaire response | 1002793 | 1002793 | |
| 1 | Daily food frequency was imputed due to missing code | 13927 | 1016720 | |
| 2 | Daily food frequency was imputed due to scanning error | 2401 | 1019121 | |
| 3 | Daily food frequency was imputed due to coding error | 0 | 1019121 | |
| < blank > | Missing | 0 | 1019121 | |