Appendix C

Statistical and Reporting Guidelines

This report presents population means and proportions, standard errors of estimates, and percentiles of dietary intake distributions. Sample weights were used to account for sample design and nonresponse. Information about the NHANES-III survey design was used in estimating variances and testing for statistical significance.

Several software packages were used to produce the tabulations:

- ***C-SIDE: Software for Intake Distribution Estimation (Version 1.0)***—used to estimate means, percentiles, and standard errors for nutrient intake tables.

- ***SUDAAN (Version 7.5)***—used to calculate means, standard errors, and tests of statistical significance for non-nutrient tables, using the DESCRIPT procedure.

- ***SAS (Version 8.2)***—used to read the NHANES-III data files, call SUDAAN procedures, process SUDAAN output, and write SUDAAN results to ASCII files.

- ***TPL (Table Producing Language)***—this software produced all data tables in appendix D.

General Procedures

NHANES-III sample weights account for the fact that each sample person does not have an equal probability of selection into the sample. NHANES-III provides sample weights for three samples: the interviewed sample weight (WTPEQX6), the MEC-examined sample weight (WTPFEX6), and the MEC and home-examined sample weight (WTPFHX6). The sampling weight used for each table in this report was specific to the data item presented in the table, and is indicated by the source of data listed in the table footnote.

Variance is generally underestimated in a complex survey when information about the survey design is not used in variance estimation. For this report, two alternate methods were used to account for the sample design.

- Balance repeated replication (BRR)—this method was specified when using C-SIDE software to obtain estimates for nutrient tables. The BRR method used the 52 replicate weights provided in the NHANES-III data.

- Taylor series linearization—this method is used in SUDAAN procedures. The complex survey design is accounted for by specifying strata and PSU in the “nest” statement of SUDAAN procedures.

Coefficients of variation (CVs) and t-statistics were generated and examined, but are not provided in the tables. CVs were examined to determine the statistical reliability of estimates, as described below in the section on Reporting Guidelines. T-statistics were examined to determine the statistical significance of differences in means and proportions. When examining categorical data, t-statistics were used and the Bonferroni adjustment was applied to adjust for multiplicity of tests.

All tests for statistical significance are tests for differences between two independent samples defined by program participation and/or income-level. In volumes I and II, differences between program participants and income-eligible nonparticipants are denoted by symbols on values for income-eligible nonparticipants; differences between program participants and higher-income nonparticipants are denoted by symbols on values for higher-income nonparticipants. In volumes III and IV, differences between the lowest-income group and the low-income group are denoted by symbols on values for the low-income group; differences between the lowest-income group and high-income group are denoted by symbols on values for the high-income group.

Differences in means and proportions were tested for statistical significance using α levels of 0.01, 0.05, and 0.001. For categorical data, differences involve multiple non-independent comparisons and were tested using α levels of 0.01, 0.05, and 0.001 adjusted using the Bonferroni method, by dividing α levels by the number of comparisons.

Age Standardization

Tables presented in appendix A include age-adjusted estimates for the total population (i.e., all age groups), calculated using the direct method (Klein, 2001). The age-adjusted estimates were obtained by weighting estimates for each age category by the year 2000 population distribution.

The population distribution used for age-adjustment is from *Monthly Estimates of the United States Population: April 2000*. Age-adjusted estimates were calculated by the SUDAAN software.

Nutrient Analyses

A primary goal for the analysis of dietary intake was to estimate the proportion of individuals whose intake is inadequate. Reference standards used to define adequate intake reflect expectations for usual intake. To apply these standards appropriately, it is necessary to have information about the distribution of intake in the population of interest. The variance of the distribution of observed intake is too large to produce reliable estimates of the prevalence of inadequate intake. This is because the variance of observed intake includes both within-person (day-to-day) and between-person variation. Methods have been established for adjusting observed intake distributions to estimate distributions of usual intake by removing within-person variation (NRC, 1986 and Nusser et al, 1996). These adjustments require two or more days of intake data for at least some subjects.

NHANES-III collected replicate 24-hour recalls on a convenience sample of approximately 5 percent of respondents. The nonrandom nature and small size of the replicate recall sample prohibited its use in estimating usual dietary intake. Instead, we used the Continuing Survey of Food Intake of Individuals (CSFII) 1994-96, to obtain estimates of within-person variation. CSFII is a nationally representative survey that includes two days of dietary intake data for all subjects.

CSFII data were used to estimate variance components for 96 demographic cells defined by age group (8), gender (male, female, both), and program participation or income (3 plus overall).[1] The variance components from CSFII were used to adjust observed intakes collected in the NHANES-III single-day dietary recalls. Estimation for all nutrients was done using *C-SIDE: Software for Intake Distribution Estimation* (Iowa State University, 1996). Because iron requirements for menstruating females are known to be asymmetrical, the adjustments performed by the C-SIDE software (using this "Iowa State Method") were not appropriate.

[1] Age groups correspond to the DRI age groups for volumes I, III, IV. CSFII used to estimate variance components for volume II (WIC participants and nonparticipants) were aggregated by year of age (4) and program participation or income (3 plus overall), but not by gender.

Therefore, distributions of iron intake were adjusted using the full probability approach as described in the IOM report *Dietary Reference Intakes: Applications in Dietary Assessment* (IOM, 2001). CSFII variance components are shown in table C1.

Reporting Guidelines

This report follows the recommendations in the NHANES-III Analytic Guidelines in the appendix titled "Joint Policy on Variance Estimation and Statistical Reporting Standards for NHANES-III and CSFII Reports: HNIS/NCHS Analytic Working Group Recommendations" (NCHS, 1996). The recommendations for presentation of statistical data call for estimates to be flagged if any of the following conditions are met:

1. **Inadequate sample size for normal approximation.** For means and for proportions based on commonly occurring events (where $0.25 < P < 0.75$), an estimate is flagged if it is based on a cell size of less than 30 times a "broadly calculated average design effect."

2. **Large coefficient of variation.** Estimates are flagged if the coefficient of variation (ratio of the standard error to the mean expressed as a percent) is greater than 30.

3. **Inadequate sample size for uncommon or very common events.** For proportions below 0.25 or above 0.75, the criteria for statistical reliability is that the cell size be sufficiently large that the minimum of nP and n(1-P) be greater than or equal to 8 times a broadly calculated average design effect, where n is the cell size and P is the estimated proportion. (I.e., an estimate is flagged when $n < 8 * (\text{avg design effect}) / \min(P,(1-P))$.) The coefficient of variation is not used in these cases.

For each data item, the design effect was calculated for each table cell as the ratio of the complex sampling design variance calculated by SUDAAN, to the simple random sample variance. The average design effect for a data item is the average of estimated design effects across age groups (pooled genders) within a demographic group, where demographic groups correspond to the columns of tables (groups defined by program participation and income).

Table C-1—CSFII variance components for 10 nutrients

Total energy

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.64563	396	0.67989	153	0.57904	646	0.64096
9-13 years old	1,160	0.60193	328	0.60372	155	0.83547	671	0.55644
14-18 years old	923	0.50309	264	0.61671	103	0.67097	549	0.40835
Male								
5-8 years old	615	0.66296	204	0.71617	82	0.59752	324	0.67722
9-13 years old	574	0.64775	150	0.66855	82	0.92401	341	0.58725
14-18 years old	474	0.56137	142	0.64129	50	0.77678	278	0.48238
Female								
5-8 years old	585	0.65811	192	0.66275	71	0.69065	322	0.64703
9-13 years old	586	0.62250	178	0.64161	73	0.86215	330	0.56842
14-18 years old	449	0.65739	122	0.73755	53	0.98718	271	0.54774

Vitamin C

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.69967	396	0.68525	153	0.59931	646	0.73627
9-13 years old	1,160	0.68751	328	0.72097	155	0.62853	671	0.69784
14-18 years old	923	0.66448	264	0.71579	103	0.64937	549	0.65834
Male								
5-8 years old	615	0.63162	204	0.65406	82	0.49264	324	0.66443
9-13 years old	574	0.75005	150	0.84172	82	0.71827	341	0.73194
14-18 years old	474	0.64366	142	0.71882	50	0.50866	278	0.64320
Female								
5-8 years old	585	0.78051	192	0.74510	71	0.74310	322	0.81309
9-13 years old	586	0.62965	178	0.64481	73	0.57962	330	0.66260
14-18 years old	449	0.71795	122	0.71471	53	0.83331	271	0.70275

Source: Variance components were estimated from two days of 24-hour recalls from the *Continuing Survey of Food Intakes by Individuals (CSFII)* using *C-SIDE: Software for Intake Distribution Estimation.*

Table C-1—CSFII variance components for 10 nutrients — Continued

Iron

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.64656	396	0.65287	153	0.50501	646	0.68303
9-13 years old	1,160	0.66462	328	0.64309	155	0.84353	671	0.63139
14-18 years old	923	0.55725	264	0.63152	103	0.63970	549	0.50132
Male								
5-8 years old	615	0.65083	204	0.65335	82	0.53149	324	0.70742
9-13 years old	574	0.70599	150	0.68009	–	–	341	0.63531
14-18 years old	474	0.64123	142	0.67109	50	0.79143	278	0.61468
Female								
5-8 years old	585	0.67081	192	0.71131	71	0.59392	322	0.69057
9-13 years old	586	0.68154	178	0.67599	73	0.70395	330	0.68327
14-18 years old	449	0.63120	–	–	53	0.75923	271	0.55397

Zinc

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.68615	396	0.70217	153	0.56120	646	0.71909
9-13 years old	1,160	0.72507	328	0.72126	155	0.93164	671	0.69166
14-18 years old	923	0.63981	264	0.70146	103	0.66696	549	0.60372
Male								
5-8 years old	615	0.71421	204	0.69299	82	0.57452	324	0.80509
9-13 years old	574	0.80163	150	0.75879	–	–	341	0.77774
14-18 years old	474	0.76450	142	0.70162	50	0.72367	278	0.81911
Female								
5-8 years old	585	0.70320	192	0.73745	71	0.66357	322	0.69698
9-13 years old	586	0.72335	178	0.77254	73	0.86709	330	0.68344
14-18 years old	449	0.71099	–	–	53	0.99831	271	0.60356

– Data not available. Estimate of within-person variance could not be obtained from CSFII.

Source: Variance components were estimated from two days of 24-hour recalls from the *Continuing Survey of Food Intakes by Individuals (CSFII)* using *C-SIDE: Software for Intake Distribution Estimation.*

Table C-1—CSFII variance components for 10 nutrients — Continued

Calcium

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.66481	396	0.69105	153	0.63871	646	0.65837
9-13 years old	1,160	0.64510	328	0.70655	155	0.63498	671	0.63231
14-18 years old	923	0.54078	264	0.59736	103	0.71750	549	0.48039
Male								
5-8 years old	615	0.67623	204	0.73051	82	0.68509	324	0.64517
9-13 years old	574	0.66883	150	0.72458	82	0.63964	341	0.66343
14-18 years old	474	0.56500	142	0.57810	50	0.66374	278	0.54808
Female								
5-8 years old	585	0.66657	192	0.65391	71	0.58929	322	0.68952
9-13 years old	586	0.65420	178	0.75460	73	0.64270	330	0.62755
14-18 years old	449	0.66246	122	0.72296	53	0.96977	271	0.57468

Total fat

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.75695	396	0.78752	153	0.69302	646	0.75267
9-13 years old	1,160	0.78766	328	0.79234	155	0.81384	671	0.78036
14-18 years old	923	0.73785	264	0.81929	103	0.84428	549	0.68569
Male								
5-8 years old	615	0.77502	204	0.77276	82	0.63507	324	0.80703
9-13 years old	574	0.84520	150	0.94216	82	0.87001	341	0.80528
14-18 years old	474	0.74780	142	0.76793	50	0.72466	278	0.76368
Female								
5-8 years old	585	0.73605	192	0.80603	71	0.74939	322	0.69713
9-13 years old	586	0.73421	178	0.68071	73	0.79022	330	0.76368
14-18 years old	449	0.72518	122	0.88103	53	0.92136	271	0.62990

Source: Variance components were estimated from two days of 24-hour recalls from the *Continuing Survey of Food Intakes by Individuals (CSFII)* using *C-SIDE: Software for Intake Distribution Estimation.*

Table C-1—CSFII variance components for 10 nutrients — Continued

Saturated fat

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.75839	396	0.75803	153	0.67629	646	0.77623
9-13 years old	1,160	0.82115	328	0.82244	155	0.88264	671	0.81170
14-18 years old	923	0.73860	264	0.84076	103	0.69077	549	0.69979
Male								
5-8 years old	615	0.77701	204	0.76687	82	0.66677	324	0.80029
9-13 years old	574	0.85686	150	0.91520	82	0.88527	341	0.82944
14-18 years old	474	0.72889	142	0.74720	50	0.63924	278	0.72752
Female								
5-8 years old	585	0.74531	192	0.75410	71	0.70250	322	0.75042
9-13 years old	586	0.78723	178	0.74058	73	0.89188	330	0.80516
14-18 years old	449	0.75136	122	0.94526	53	0.72077	271	0.69149

Cholesterol

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.71983	396	0.79130	153	0.73983	646	0.71434
9-13 years old	1,160	0.77319	328	0.82575	155	0.85483	671	0.74862
14-18 years old	923	0.76702	264	0.86552	103	0.84932	549	0.68532
Male								
5-8 years old	615	0.69441	204	0.81756	82	0.62241	324	0.67669
9-13 years old	574	0.82410	–	–	82	0.81907	341	0.82416
14-18 years old	474	0.79341	–	–	50	0.74427	278	0.73885
Female								
5-8 years old	585	0.76591	192	0.77679	71	0.89985	322	0.77100
9-13 years old	586	0.76725	178	0.80748	73	0.92618	330	0.73228
14-18 years old	449	0.82269	122	0.80299	–	–	271	0.77642

– Data not available. Estimate of within-person variance could not be obtained from CSFII.

Source: Variance components were estimated from two days of 24-hour recalls from the *Continuing Survey of Food Intakes by Individuals (CSFII)* using *C-SIDE: Software for Intake Distribution Estimation.*

Table C-1—CSFII variance components for 10 nutrients — Continued

Sodium

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.68844	396	0.65403	153	0.65134	646	0.72932
9-13 years old	1,160	0.70650	328	0.75644	155	0.88986	671	0.64068
14-18 years old	923	0.65402	264	0.79165	103	0.74218	549	0.55714
Male								
5-8 years old	615	0.66953	204	0.66852	82	0.64605	324	0.73027
9-13 years old	574	0.72338	150	0.83047	82	0.96578	341	0.61501
14-18 years old	474	0.70948	142	0.79749	50	0.78070	278	0.65162
Female								
5-8 years old	585	0.74113	192	0.66840	71	0.77560	322	0.76133
9-13 years old	586	0.75399	178	0.77111	–	–	330	0.73251
14-18 years old	449	0.79052	122	0.90254	–	–	271	0.68420

Fiber

	All children		Lowest income: ≤ 130% poverty		Low-income: 131-185% poverty		Higher-income: > 185% poverty	
	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance	Sample size	Within-individual variance
Both sexes								
5-8 years old	1,200	0.72267	396	0.78046	153	0.64452	646	0.69572
9-13 years old	1,160	0.67375	328	0.67523	155	0.58027	671	0.70963
14-18 years old	923	0.67390	264	0.76122	103	0.76593	549	0.61061
Male								
5-8 years old	615	0.76335	204	0.78724	82	0.71858	324	0.75792
9-13 years old	574	0.69010	150	0.74114	82	0.58389	341	0.69763
14-18 years old	474	0.72004	142	0.77868	50	0.84791	278	0.65561
Female								
5-8 years old	585	0.68392	192	0.78455	71	0.61234	322	0.65287
9-13 years old	586	0.68649	178	0.68586	73	0.65453	330	0.73565
14-18 years old	449	0.72705	122	0.77933	53	0.90248	271	0.66310

– Data not available. Estimate of within-person variance could not be obtained from CSFII.

Source: Variance components were estimated from two days of 24-hour recalls from the *Continuing Survey of Food Intakes by Individuals (CSFII)* using *C-SIDE: Software for Intake Distribution Estimation.*