Introduction

- Background
- Addressing Data Needs
- Public Health Uses
- Data Presented for Each Biochemical Indicator
- Interpreting the Data
- Useful Sources of Information about Using Nutrition Monitoring to Interpret Data
- The National Health and Nutrition Examination Survey (NHANES)
- Data Analysis
- References

Introduction

Background

The National Report on Biochemical Indicators of Diet and Nutrition in the U.S. Population is a series of publications that provide ongoing assessment of the U.S. population's nutritional status by measuring blood or urine concentrations of diet-and-nutrition biochemical indicators. The Centers for Disease Control and Prevention's (CDC) Division of Laboratory Sciences at the National Center for Environmental Health (NCEH/DLS) conducted the laboratory analyses for 58 biochemical indicators presented in this 2012 report, which is the second in this series. CDC measured these indicators in specimens from a representative sample of the U.S. population during all or part of the four-year period from 2003 through 2006. Where available, data are also presented on changes of biochemical indicator concentrations over time since 1999. Similarly, data are also presented on the prevalence of low or high biochemical indicator concentrations during 2003–2006, and on changes in the prevalence over time since 1999. The first report of this series was published in July 2008 and contains information on 27 biochemical indicators from all or part of the four-year period from 1999 through 2002. Both reports can be accessed online: http://www.cdc.gov/nutritionreport.

Characteristic	First report, 2008	Second report, 2012
Years of NHANES covered	1999–2002	2003–2006
Number of indicators covered	27	58
Concentrations by race/ethnic group	Yes	Yes
Central 95% reference intervals	No	Yes
Graphic representation of age patterns	No	Yes
Concentrations over time	No	Yes (1999-2006)
Prevalence estimates	No	Yes
Prevalence estimates over time	No	Yes (1999-2006)

CDC's National Health and Nutrition Examination Survey (NHANES), conducted by the National Center for Health Statistics (NCHS), collected the specimens for this report. NHANES is a series of surveys designed to collect data on the health and nutritional status of the U.S. population. This report covers biochemical measurements—one important facet in the assessment of the U.S. population's nutritional status. Other nutrition-related aspects from NHANES, such as dietary intake, supplement usage, hematologic measurements, and anthropometric body measurements are not covered.

In this report, a biochemical indicator means a nutrient (e.g., vitamin, fatty acid, trace element), a metabolite (e.g., homocysteine, methylmalonic acid), or a dietary indicator with potential health relevance (e.g., isoflavone, lignan) measured in blood or urine. Although most biochemical indicators presented in this report enter the human body from foods or supplements, the body itself produces some indicators in response to dietary intake or environmental exposure. Blood and urine concentrations reflect the amount of nutrients and dietary compounds actually in the body or passing through the body from all these sources.

The biochemical indicator sections and new biochemical indicators covered in this report are:

Biochemical indicator sections
• Water-soluble vitamins
• Fat-soluble vitamins and nutrients
• Trace elements (iron indicators and iodine)
• Isoflavones and lignans
• Acrylamide hemoglobin adducts

New biochemical indicators
• Vitamin B6
• Vitamin C
• Fatty acids
• Iron status: Transferrin receptor and body iron
• Acrylamide hemoglobin adducts

Addressing Data Needs

This report is the second CDC product containing reference information on NCEH/DLS measurement data for a wide range of biochemical indicators of diet and nutrition from the most recent continuous NHANES survey, starting in 1999. In this comprehensive report, information on changes in concentrations of a large number of biochemical indicators during 1999–2006 is presented for the first time. Prevalence information on low or high biochemical indicator concentrations is also presented for the first time.

NCHS has historically released or commissioned a variety of products presenting NHANES results. Among these are Data Briefs, Data Tables, Advance Data, Series Reports, and Reports through the Life Sciences Research Office (LSRO). NHANES Series Reports (mainly Series 11) and LSRO Reports from surveys prior to the continuous NHANES have been of particular value to the nutrition community (see **Appendix A**). The NHANES Web site provides current information on results and products from this survey: http://www.cdc.gov/nchs/about/major/nhanes/survey_results_and_products.htm.

Public Health Uses

This report's primary objective is to inform public health scientists and policy makers about the concentrations of biochemical indicators of diet and nutrition in the general U.S. population and in selected subpopulations. These data will help physicians, scientists, and public health offcials assess inadequate or excess intake and will inform analyses on the relation between biochemical indicators and health outcomes. Other objectives and potential public health uses of the information include

- Establishing and improving on existing population reference levels that can be used to determine whether an individual or a group has an unusually high or low concentration of a diet-and-nutrition biochemical indicator.
- Determining whether the nutritional status of special population groups, such as minorities, children, women of childbearing age, or the elderly, is different from that of other groups, or whether such nutritional status needs improvement.
- Tracking trends over time in the population's biochemical indicator concentrations.
- Assessing the effectiveness of public health efforts to improve the diet and nutritional status of the U.S. population.
- Guide research to perform more in-depth analyses of the NHANES data and to generate hypotheses for future nutrition and human health studies.

Data Presented for Each Biochemical Indicator

This report contains tables and figures of descriptive statistics on the distribution of blood and urine concentrations during all or part of the four-year period from 2003 through 2006 for each diet-and-nutrition biochemical indicator. Statistics include unadjusted geometric means and selected percentiles with confidence intervals. For some biochemical indicators, additional information is included, as available, in the form of

- Tables and figures describing biochemical indicator concentrations across survey cycles during all or part of the eight-year period from 1999 through 2006. Statistics include unadjusted geometric means and selected percentiles with confidence intervals.
- Tables describing the prevalence of low or high concentrations of selected biochemical indicators during all or part of the four-year period from 2003 through 2006 and tables describing the prevalence across survey cycles during all or part of the eight-year period from 1999 through 2006. Statistics include unadjusted percentages with confidence intervals and estimated total number of persons affected.

See **Appendix B** for an overview of the type of information presented for each biochemical indicator. The data are grouped by age, gender, and race/ethnicity. The majority of the biochemical indicators reviewed in this report, with the exception of vitamin C and body iron, have a long upper tail (skewed right). For these biochemical indicators, a geometric mean provides a better estimate of central tendency because it is less influenced by high values than is the arithmetic mean. However, the arithmetic mean is presented for vitamin C and body iron as the distributions for these biochemical indicators were reasonably symmetric. Scientists can use the presented percentile levels to determine those serum, blood, or urine indicator concentrations common to people in the U.S. population and those that are unusual. Frequently, the central 95% reference interval (2.5th to 97.5th percentile) is used to describe normal concentrations in a population. Concentrations outside the reference interval are considered unusual. For urine measurements, data are shown for both the concentration and for the concentration corrected for the urinary creatinine level.

We present the following information for each biochemical indicator during all or part of 2003–2006:

- A table that presents the geometric mean and selected percentile (2.5th–97.5th, so called central 95% reference interval) concentrations by age, gender, or race/ethnicity (1-level stratified).
- A figure that presents the geometric mean concentrations by age and gender or by age and race/ethnicity (2-level stratified).
- Four detailed tables that present the geometric mean and selected percentile (5th or 10th, 50th, 90th or 95th) concentrations by age, gender, and race/ethnicity (3-level stratified). The first table is for the overall U.S. population stratified by age and gender, while the next three tables present data for each racial/ethnic group (Mexican American, non-Hispanic black, and non-Hispanic white) stratified by age and gender.

If data are available for multiple two-year survey cycles from 1999–2006, we present tables with geometric mean and selected percentile (5th–95th) concentrations by age, gender, or race/ethnicity for each available two-year survey cycle, as well as corresponding figures for selected percentiles (1-level stratified).

For biochemical indicators that have accepted cutoff values for low or high concentrations or for both (e.g., folate, vitamins A, B6, B12, C, D, E, ferritin, iodine)—suggesting deficiency or excess of certain micronutrients—we present tables with prevalence estimates by age, gender, or race/ethnicity during all or part of 2003–2006 (1-level stratified). We also present tables with prevalence estimates by age, gender, or race/ethnicity for each available two-year survey cycle from 1999–2006 to allow evaluation of changes in prevalence estimates over time (1-level stratified). See **Appendix C** for a complete listing of the cutoff values and populations described in this report.

Background text provides general information for each indicator to aid with interpreting the data.

To address sources of these nutrients, biochemical pathways in the body, and known health effects, the text contains a brief overview about each indicator.

Each chapter contains highlights followed by detailed observations that are derived from this report's data tables and figures.

The *highlights* are presented directly after the background text. They summarize important observations and discuss them in a public health context. For example, we present figures highlighting prevalence information by demographic subgroups. Where long-term trending information beyond the continuous NHANES is available and of public health interest, we present figures showing changes in biochemical indicator concentrations from NHANES III (1988–1994) to NHANES 1999–2002 and NHANES 2003–2006. The *detailed observations* describe selected categorical differences between demographic subgroups derived from the data tables and figures that follow next. Each chapter is concluded by a list of pertinent references.

Interpreting the Data

Blood or urine concentrations of biochemical indicators can help in assessing the adequacy of intake for the U.S. population. These measurements indicate cumulative intakes from foods, some fortified with micronutrients (e.g., iron, thiamin, riboflavin, niacin, folate, vitamin A, vitamin D), and from dietary supplements that contain vitamins, minerals, or both. However, blood or urine concentrations of biochemical indicators can also be influenced by factors other than diet, such as various diseases or exposures. For those nutrients without defined adequate intakes (e.g., carotenoids, isoflavones), biochemical indicators are useful for assessing intake without regard to adequacy.

Dietary deficiencies are well documented, and they have characteristic signs and symptoms. In addition, recent findings have determined that less than optimal biochemical concentrations (representing suboptimal status) have been associated with risks of adverse health effects. These health effects include cardiovascular disease, stroke, impaired cognitive function, cancer, eye diseases, poor bone health, and other conditions. Adverse health effects, including toxicity, are also possible from consuming excess amounts of certain nutrients and data to assist in the assessment of excessive intake is a feature of this report. Determining the concentrations of a biochemical indicator that may indicate risk for disease and the concentrations that are of negligible health concern requires future research studies that are separate from this report. In collaboration with other agencies and institutions, CDC encourages, and itself conducts research on the relationship between biochemical indicators and health effects.

This report contains unadjusted geometric means, selected percentiles, and prevalence estimates of low or high concentrations of diet-and-nutrition biochemical indicators for the civilian, noninstitutionalized U.S. population. A limited interpretation of relative differences between population groups is possible by identifying groups with nonoverlapping confidence intervals. However, one should be careful about interpreting the observed differences as causal. The intent is to describe the characteristics of the population and of selected subgroups, not to explain why the groups display certain characteristics or why they differ from each other. Furthermore, differences in biochemical indicator concentrations of selected subgroups do not necessarily imply health status problems. And for several reasons, one should use caution when drawing temporal conclusions from comparisons of serial cross-sectional NHANES survey cycles. One of these reasons is that different or improved methods of measurement may be employed across the NHANES survey cycles. Another reason is that there are demographic changes to the U.S. population over time. Finally, sampling differences could explain some of the observed changes from one cycle to the next. More in-depth statistical analyses, such as developing models to adjust simultaneously for many covariates and taking into consideration interactions between two or more variables, are beyond the scope of this report. Nonetheless, unadjusted geometric means, selected percentiles, and prevalence estimates provided in this report are useful to summarize reference information for blood or urine concentrations of diet-and-nutrition biochemical indicators for the civilian, noninstitutionalized population in the United States and selected subgroups. We hope that the report will stimulate scientists to examine the data further through analyzing the raw data available at: http://www.cdc.gov/nchs/nhanes.htm.

Laboratories may use different methods for measuring the indicators reported here. However, different methods may result in different method-specific reference intervals. Consequently, to apply these results, health science professionals should check with their particular laboratory to be sure that their methods compare closely to those used in this report (see **Appendix D**).

Sources of Information on Nutrition Monitoring to Help Interpret the Data

Information about dietary intake is critical to research examining the reasons for nutritional inadequacies. Such information is also critical to programs seeking to improve diet and nutritional status. Selected NCHS Advance Data Reports provide useful overviews (see **Appendix A**). Also of value are the U.S. Department of Agriculture's (USDA) databases on food surveys and food composition:

What We Eat in America (WWEIA) is the dietary intake interview section of NHANES (http://www.ars.usda.gov/foodsurvey).

The Food and Nutrient Database for Dietary Studies (FNDDS) (http://www.ars.usda.gov/Services/docs.htm?docid=12089) is a database of foods, their nutrient values, and weights for typical food portions. This database is used to generate data for the WWEIA survey through application of the nutrient values from the National Nutrient Database for Standard Reference (http://www.ars.usda.gov/Services/docs.htm?docid=8964).

The National Health and Nutrition Examination Survey (NHANES)

CDC laboratory scientists used biological specimens obtained from NHANES participants to measure biochemical indicators of diet and nutrition for this publication. NHANES is a series of NCHS-conducted surveys designed to collect data on the health and nutritional status of the U.S. population. This is the only national survey that collects biological samples. The NHANES surveys began in 1960 with the first Health Examination Survey (HES 1). The nutritional component was added in the early 1970s in NHANES I. In 1999, NHANES became a continuous survey, sampling the U.S. population annually and releasing the data in two-year cycles.

NHANES collects information on a wide range of health-related behaviors, conducts physical examinations, and collects samples for laboratory tests. Because of physical examination and biological measures, NHANES is unique in its ability to examine public health issues in the U.S. population, such as risk factors for cardiovascular disease. To select a representative sample of the civilian, noninstitutionalized population in the United States, the survey sampling plan follows a complex, stratified, multistage, probability-cluster design. The civilian, noninstitutionalized population consists of persons who are neither in the military nor institutionalized (e.g., they are not residents of nursing homes, college dormitories, or prisons).

The NHANES protocol includes a home interview followed by a standardized physical examination at a mobile examination center. As part of the examination, for participants aged 1 year and older, blood is obtained by venipuncture. Urine specimens are collected from participants aged 6 years and older. By design, approximately half of the participants are evaluated after an overnight fast; for the other half of the participants, there is approximately an equal distribution between those who fasted less than 3 hours and those who fasted between 3 and 8 hours before providing a biological sample. Because weather can adversely affect the mobile examination centers, data are collected in northern latitudes in summer and in southern latitudes in winter. This seasonal-latitude structure might indirectly affect biochemical indicators.

Additional detailed information about the design and conduct of the NHANES survey is available at http://www.cdc.gov/nchs/nhanes.htm. Information about how biological specimens are collected is available at (http://www.cdc.gov/nchs/data/nhanes/blood.pdf) and included in the Laboratory Procedures Manual at http://www.cdc.gov/nchs/data/nhanes/lab1-6.pdf and at http://www.cdc.gov/nchs/data/nhanes/lab7-11.pdf.

Data Analysis

NCHS has developed a comprehensive Web-based tutorial (http://www.cdc.gov/nchs/tutorials/Nhanes/index.htm) to help users better understand the complex survey design and to help them analyze NHANES data.

Because the NHANES sample design is a complex, multistage probability sample, officials use sample weights when estimating the mean or other descriptive metrics. These weights are post-stratified to the U.S. Census Bureau estimates of the U.S. population to adjust for the unequal probability of selection into the survey and possible bias resulting from nonresponse. Demographic data files released by NCHS for each NHANES two-year survey cycle include a two-year interview weight and a medical examination weight. All estimates in this report use the appropriate medical examination weight. The selected medical examination weight depends on whether the specimens tested constitute a random subsample of all the eligible participants and how many survey cycles are combined to produce the estimate.

Combining data over multiple survey cycles can produce estimates with increased statistical reliability. In cases of combined estimates, new weights were constructed. For example, a four-year estimate for the years 2003–2006 was based on a four-year weight, which was created by assigning half the two-year weight for 2003–2004 or half the two-year weight for 2005–2006, depending when the person was sampled.

Results are shown for the total population and by age group, gender, and race/ethnicity, as defined in NHANES. For these analyses, race/ethnicity is presented as Mexican American, non-Hispanic black, and non-Hispanic white. Other racial or ethnic groups are sampled, but the proportion of the total population represented by these other groups is not large enough to produce valid estimates. Thus, this report does not include separate estimates for other racial subcategories. The other racial/ethnic groups, however, are included in the overall estimates.

Data were analyzed by use of the statistical software package Statistical Analysis System (SAS, Version 9.2) and SUDAAN (Release 10.0). SUDAAN uses sample weights and calculates variance estimates that account for the complex survey design. Guidelines for the analysis of NHANES data are provided by NCHS at: http://www.cdc.gov/nchs/nhanes/nhanes2003-2004/analytical_guidelines.htm.

Standard error estimates were calculated by use of the Taylor series (linearization) method within SUDAAN. The degrees of freedom for variance estimation are generated by subtracting the number of strata from the number of primary sampling units (PSUs).

Geometric means were calculated by taking the log of each concentration, calculating the mean of those log values, then taking the antilog of that mean (the calculation can be done by use of any log base, such as 10 or e). The confidence interval uses the standard error and mean on the log scale and the appropriate critical value from the t-distribution to calculate upper and lower confidence limits on the log scale. The confidence intervals of geometric means in this report are based on taking the antilog of those upper and lower confidence intervals.

Percentile estimates were calculated by use of linear interpolation. Confidence intervals for percentiles were calculated by the Woodruff method (1952). This method uses the standard error of the empirical distribution function at the selected percentile and constructs a 95% confidence interval, followed by back-transforming by use of the inverse of the empirical distribution (see **Appendix E** for more details). We used the unweighted sample size and assumed an average design effect of 1.4 as the criteria to estimate percentiles of sufficient precision (U.S. Centers for Disease for Control and Prevention 1996; Table 1 in Appendix B). In order for percentiles to be considered reliable, at least 112 persons had to be represented to allow estimation of the 10th and 90th percentiles, 224 persons for the 5th and 95th percentiles, and 448 persons for the 2.5th and 97.5th percentiles. We flagged and footnoted percentiles where these requirements were not met.

Prevalence estimates for low or high concentrations of biochemical indicators are the weighted percentage of persons who fall below or above a predefined cutoff value (see **Appendix C**). The confidence intervals for prevalence estimates are based on a logit transformation that ensures the confidence interval limits cannot fall outside of 0 and 1. We used the relative standard error (RSE) as a criterion for prevalence estimates of sufficient precision. The RSE is calculated as a percentage by dividing the standard error of the estimate by the estimate value and multiplying by 100. Prevalence estimates associated with a RSE between 30% and less than 40% are flagged and footnoted in this report. Estimates are not provided if they are associated with an RSE equal to or greater than 40%.

Estimates of the total number of persons who met the definition of having low or high concentrations of biochemical indicators were generated by multiplying the weighted prevalence estimate by the population size of interest, derived from the current population survey (CPS) at the midpoint of the available two-year cycle. Confidence intervals for the estimated total, while not presented, are calculated by multiplying the population size of interest by the upper and lower limits of the 95% confidence interval for the weighted prevalence. CPS-based population tables for NHANES by age, gender, and race/ethnicity are on the NHANES Web page for a given survey cycle, available at http://www.cdc.gov/nchs/nhanes/response_rates_CPS.htm. When estimates of the total count are based on combined survey cycles (2003–2006), the 2003–2004 CPS-based population table at the above link was used.

Figures in the highlight section that present age-adjusted geometric mean concentrations from NHANES III (1988–1994), NHANES 1999–2002, and NHANES 2003–2006 or age-adjusted prevalence estimates by demographic subgroups have been generated in SUDAAN by use of age-standardizing proportions from the 2000 U.S. Census population (using direct standardization). Statistically significant differences between age-adjusted geometric means and age-adjusted prevalences were assessed through pairwise comparisons. A reader should take care when interpreting these age-adjusted figures in isolation. The magnitude of an age-adjusted geometric mean or age-adjusted prevalence is completely arbitrary, and it depends upon the chosen standard population. Additionally, age-adjusted geometric means or age-adjusted prevalences can mask important information about trends if age-specific rates do not have a consistent relationship. It is worth noting that while NHANES 1999–2004 provided a race/ethnicity variable that was an analytic link to the NHANES III race/ethnicity variable called RIDRETH2, this variable is not included in the NHANES 2005–2006 demographics file. Therefore, the codes of the race/ethnicity variable called RIDRETH1 were used in displaying age-adjusted geometric means by race/ethnicity for NHANES 1999–2002 and NHANES 2003–2006. RIDRETH1 includes all multi-racial responses in the Other category; whereas, RIDRETH2 includes multi-racial responses for Non-Hispanics with primary race White or Black in the Non-Hispanic White or Non-Hispanic Black categories. This means that there are slightly fewer people coded as non-Hispanic white and non-Hispanic black through RIDRETH1 than for RIDRETH2; however, this difference does not affect the Mexican American category.

The limit of detection (LOD) is the level at which the measurement has a 95% probability of being greater than zero (Taylor 1987). For calculation of geometric means, concentrations less than the LOD were assigned a value equal to the LOD divided by the square root of 2. If the proportion of results less than the LOD (< LOD) was greater than 40%, geometric means were not calculated. Percentile estimates less than the LOD were reported as "< LOD". Most of the indicators had very few results below the LOD value, so that the choice of statistical analysis to handle these results makes little practical difference. There were a few exceptions, however (e.g., serum *cis-beta*-carotene, retinyl palmitate and retinyl stearate; urinary O-desmethylangolensin and equol), where a larger proportion of results were < LOD. **Appendix F** contains a table of LOD values for each biochemical indicator, as well as the unweighted percent of data values that were < LOD for each survey cycle. LOD values may change over the time period of the report as a result of changes in analytical methods. We used the higher of the two LOD values for the analysis of the combined four-year data for 2003–2006.

Due to changes to analytical methods for plasma total homocysteine, serum 25-hydroxyvitamin D, and serum ferritin, an adjustment equation was applied to the data, as described in the NHANES documentation:

- http://www.cdc.gov/nchs/nhanes/nhanes1999-2000/LAB06.htm#Component_Description
- http://www.cdc.gov/nchs/nhanes/nhanes2003-2004/L06VID_C.htm
- http://www.cdc.gov/nchs/nhanes/nhanes2003-2004/L06TFR_C.htm#Analytic_Notes.

For biochemical indicators measured in urine, we present separate tables for the concentration of the indicator expressed as "per volume of urine" (uncorrected table) and the concentration of the indicator expressed as "per gram of creatinine" (creatinine-corrected table). Comparison of an individual participant's result to population data in the tables requires correction for urinary dilution; thus, an individual creatinine-corrected result is needed and should be compared to the creatinine-corrected data tables. Otherwise, health scientists may compare means and percentiles from other studies to the tables having either of the corresponding units. We used the uncorrected tables to compare urine concentrations across groups. Because instrument responses are measured in units of weight per volume, LOD calculations were performed by use of the concentration of the indicator expressed as per volume of urine. For this reason, LOD results for urine measurements in **Appendix F** are in weight per volume of urine. In the creatinine-corrected tables, a result for a geometric mean or a percentile was reported as < LOD if the corresponding geometric mean or percentile was < LOD in the uncorrected table. Thus, for example, if the 5th percentile for males was < LOD in the uncorrected table, it would also be < LOD in the creatinine-corrected table.

References

Taylor JK. Quality assurance of chemical measurements. Chelsea (MI): Lewis Publishing; 1987.

U.S. Centers for Disease Control and Prevention. NHANES analytic guidelines, the Third National Health and Nutrition Examination Survey, NHANES III (1988–94). Hyattsville (MD): National Center for Health Statistics; October 1996 [cited 2011]. Available at: http://www.cdc.gov/nchs/data/nhanes/nhanes3/nh3gui.pdf.

Woodruff RS. Confidence intervals for medians and other position measures. J Am Stat Assoc. 1952;57:622–627.