



Number 24

Statistical Notes



July 2002

Healthy People 2010 Criteria for Data Suppression

Richard J. Klein, M.P.H.; Suzanne E. Proctor, M.S.P.H.; Manon A. Boudreault, M.P.H.;
and Kathleen M. Turczyn, M.P.H.

Introduction

Healthy People 2010 is the third generation of 10-year goals and objectives for the Nation, and builds on initiatives pursued over the past two decades (1–3). It identifies 467 objectives for improving the health of the Nation. Because many of these objectives include a number of subobjectives, there are a total of 823 separate measures presented in *Healthy People 2010*. Of these measures, 432 are population-based.

To facilitate comparisons among population groups, the Healthy People 2010 objectives include a population data table or “template” for each population-based measure. This template includes, at a minimum, disaggregations for specific race, Hispanic origin, gender, and socioeconomic status (either education or income). In addition, some objectives show further disaggregations for categories such as age, disability status, urban/rural classification, specific health conditions, or sexual orientation (3,4).

In *Healthy People 2010* baseline data are shown for each data cell in the population template, where available. If data are not available, the cell contains one of the following abbreviations:

- DNC—Data are not collected by the data system used to monitor the objective.
- DNA—Data have been collected but have not yet been analyzed.
- DSU—Data do not meet the criteria for statistical reliability, data quality, or confidentiality (data are suppressed).

The first two categories, DNC and DNA, are self-explanatory. However, there are a number of reasons related to statistical reliability, data quality, or confidentiality why data are suppressed (DSU). In general, there are four major reasons a statistic in *Healthy People 2010* is suppressed:

Acknowledgments

We are indebted to the following individuals for their contributions to the Data Systems section: Joyce Abma, CDC, NCHS (NSFG); Lawrence Agodoa, NIH, NIDDK (USRDS); Anjani Chandra, CDC, NCHS (NSFG); Donald Cherry, CDC, NCHS (NAMCS and NHAMCS); Randy Curtin, CDC, NCHS (NHANES); Brady Hamilton, CDC, NCHS (NVSS-N); Deborah Holtzman, CDC, NCCDPHP (BRFSS); Joe Goldman, USDA (CSFII); Ruth Ann Jajosky, CDC, EPO (HIV/AIDS, NNDSS, and STDSS); Lloyd Johnston, University of Michigan (MTF); Laura Kann, CDC, NCCDPHP (YRBSS and SHPPS); Carol Knowles, CDC, EPO (NNDSS); Kenneth Kochanek, CDC, NCHS (NVSS-M); Doris Lefkowitz, AHRQ (MEPS); Wendy Lynch, AWHP (NWHP); Maria Owings, CDC, NCHS (NHDS); Michael Rand, DOJ, BJS (NCVS); Susan Schober, CDC, NCHS (NHANES); Patricia Sweeney, CDC, NCHSTP (HIV/AIDS); and David Woodwell, CDC, NCHS (NAMCS and NHAMCS). For their thoughtful review of the manuscript, we also thank Jennifer Madans, Associate Director for Science, CDC, NCHS; Carter Blakey and Diane Wagener of the Office of Disease Prevention and Health Promotion, DHHS; and Van Parsons of the Office of Research and Methodology, CDC, NCHS. This report was edited by Thelma W. Sanders and typeset by Jacqueline M. Davis of the Publications Branch, Division of Data Services.

- Number of events is too small to produce reliable estimates or may violate confidentiality requirements.
- Sample design does not produce representative estimates for a particular group.
- High item nonresponse or large number of unknown entries.

This report discusses the specific criteria for data suppression used by 22 of the 23 major data systems used to track Healthy People 2010 objectives. A major data system is defined as one that is used to track five or more objectives. The State Tobacco Activities Tracking and Evaluation System (STATES) is not discussed, as it is an administrative database that produces complete counts of States involved in various activities.

Most of the major data systems discussed below are based on sample surveys (for example, the National Health Interview Survey or the National Household Survey on Drug Abuse), thus the criteria for statistical reliability generally involve consideration of sampling error. Several systems rely on population counts of events (for example, the National Vital Statistics System or the National Notifiable Diseases Surveillance System). Although data from these systems are not subject to sampling error, when used for analytic purposes these data may be subject to random variation, especially when the number of events is small. Issues of confidentiality may also restrict the presentation of data from these systems. Bias resulting from data that are not representative of the population or where there is high item nonresponse is a concern for sample surveys and population data systems.

Different criteria for data suppression have been adopted by the data systems used to monitor the health status of the U.S. population. Details on the criteria for suppressing data (data cell shown as DSU) from the major Healthy People 2010 data systems are discussed in the following sections. In general, the Healthy People 2010 objectives are tracked with national data. However, because the objectives are also of interest at the State and local level, information on suppression of subnational data is also provided where applicable.

Table 1 shows a summary of the criteria for suppression as well as information on the specific population groups for which data systems are not able to show reliable estimates (that is, population groups always displayed as DSU in Healthy People 2010) for each of the Healthy People 2010 major data systems. These are primarily race and/or Hispanic origin groups.

Data Systems

The information below represents general guidelines established by the sponsoring agencies for presenting data from their data systems for tracking the Healthy People 2010 objectives. It is possible that different criteria may be used in other publications and electronic data presentations. The criteria described below apply to the specific measures used to track the Healthy People 2010 objectives, such as

proportions, crude rates/ratios, and age-adjusted rates/ratios (5,6). These measures are defined in the Healthy People 2010 statistical compendium, *Tracking Healthy People 2010* (4). Additional information on the sponsoring agencies and Internet addresses for these surveys is shown in table 2.

Sample surveys

Behavioral Risk Factor Surveillance System (BRFSS)

BRFSS is an ongoing, State-based system of health surveys conducted by telephone interview using random-digit-dialed probability samples of adults ages 18 years and over (7,8). The national sample size, which is increasing annually, yielded over 200,000 interviews in 2001. Major topics include alcohol and tobacco use, physical activity, nutrition, screening for certain cancers, health status, health care access, hypertension awareness, and diabetes awareness. Some topics are asked annually, some biennially, and others at the discretion of individual States. Estimates are considered statistically unreliable and are suppressed if the denominator is based on fewer than 50 sample cases.

Continuing Survey of Food Intake by Individuals (CSFII)

CSFII is a nationally representative periodic survey, which uses a stratified multistage probability sample of the U.S. noninstitutionalized civilian population (9,10). For 1994–96, the sample size was about 16,000 individuals of all ages, and for 1998, an additional sample was collected from about 5,500 children aged 0–9 years. CSFII includes the collection of data on the kinds and amounts of foods consumed on each of two nonconsecutive days, sources of foods, time, and name of each eating occasion. Information on food expenditures, shopping practices, pregnancy, lactation, nursing status, and height and weight are also collected. United States Department of Agriculture (USDA) follows the recommendations on statistical reporting developed jointly by USDA and DHHS for CSFII and NHANES III reports (11). Estimates between 25 percent and 75 percent are considered unreliable and are suppressed if the coefficient of variation of the mean (CV) is greater than 30 percent or the sample size is less than 30 times an average or generalized design effect. A variance inflation factor (VIF), equal to the ratio of the mean of the squared sampling weights to the square of the mean of the weights, is used in the generalized design. Estimates of 25 percent or less are considered unreliable and are suppressed if the sample size is less than eight times a generalized design effect divided by p , where p is the proportion expressed as a fraction. Estimates of 75 percent or greater are considered unreliable and are suppressed if the sample size is less than eight times a generalized design effect divided by $(1-p)$. For all objectives tracked by this survey, data for American Indians or Alaska Natives and Asians or Pacific Islanders are suppressed because the size of the sampled populations for these groups is too small. CSFII has been integrated with the National Health and Nutrition Examination Survey (NHANES); dietary data collection for the integrated survey

began in 2002. More information on statistical reliability for data from this survey is available in a report by the Human Nutrition Information Service and the National Center for Health Statistics (12).

Medical Expenditure Panel Survey (MEPS)

MEPS is an annual, nationally representative subsample of respondents to the National Health Interview Survey (NHIS) and uses the stratified, multistage probability sample design of NHIS. The sample size is about 25,000 individuals. The survey collects data on the civilian noninstitutionalized population with four linked, integrated surveys using computer-assisted personal interviews, telephone interviews, and mailed surveys (13). General survey content includes health conditions, health status, use of medical care service, charges and payments, and health insurance coverage and plans. Estimates are considered statistically unreliable and are suppressed if 1) the denominator is based on fewer than 70 sample cases, or 2) the relative standard error (RSE) of the estimate (expressed as a percent) is greater than 30 percent. Healthy People 2010 tracking data for American Indians or Alaska Natives and Asians or Pacific Islanders are suppressed because of their small numbers in the survey.

Monitoring the Future (MTF)

The MTF study uses self-administered questionnaires in annual national surveys of representative samples of 8th, 10th, and 12th graders in public and private schools in the continental United States (14,15). The sample size is about 45,000–50,000 students. General survey content includes a wide range of behaviors, attitudes, and values, but drug use and related attitudes are among the key variables. There are no formal criteria for statistical reliability. In general, data may be considered statistically unreliable when there are fewer than 100 cases in the denominator or if data values are less than 0.05 percent. For all objectives tracked by this survey, data for American Indians or Alaska Natives and Asians are statistically unreliable and are suppressed. There are not enough persons sampled from these population groups to produce reliable estimates.

National Ambulatory Medical Care Survey (NAMCS) and National Hospital Ambulatory Medical Care Survey (NHAMCS)

The annual NAMCS and NHAMCS use multistage probability sample designs. NAMCS collects data from over 3,000 nonfederally employed office-based physicians and NHAMCS collects data on office visits from about 400 emergency departments and about 230 outpatient departments on samples of their patient visits during an assigned reporting period (16,17). Key variables include physician practice characteristics, patient characteristics, and numerous visit characteristics such as reason for visit, physician's diagnosis, and visit disposition. Estimates are considered statistically unreliable if 1) the numerator is less than 30, or 2) the RSE of the estimate is greater than 30 percent. For all objectives tracked by these surveys, data

for American Indians or Alaska Natives, Asians or Pacific Islanders, and Hispanics are suppressed. For American Indians or Alaska Natives and Asians or Pacific Islanders, the number of visits is too small. For the Hispanic origin variable, item nonresponse is too high. In 1998 Hispanic origin was missing for 20 percent of the records; in 1999, it was missing for 22 percent.

National Crime Victimization Survey (NCVS)

NCVS is an annual, nationally representative survey of the civilian, noninstitutionalized population ages 12 years and over. A stratified, multistage cluster sampling strategy is used to select approximately 50,000 households for a series of telephone and in-person interviews (18,19). The survey obtains information on the frequency, nature, and consequences of criminal victimizations (including those not reported to police). Estimates are considered statistically unreliable and are suppressed if they are based on 10 or fewer sample cases in the numerator.

National Health Interview Survey (NHIS)

NHIS uses a stratified, multistage probability sample design. It collects data annually on the civilian, noninstitutionalized population by computer-assisted personal interview (20,21). The expected sample of 43,000 occupied respondent households yields a probability sample of about 111,000 persons. General survey content includes chronic health conditions and limitations in activity, health behaviors, health care access, health care provider contacts, immunizations, and acquired immunodeficiency syndrome (AIDS) knowledge and attitudes. Estimates are considered statistically unreliable and are suppressed if 1) the denominator is based on fewer than 50 sample cases, or 2) the RSE of the estimate is greater than 30 percent.

National Health and Nutrition Examination Survey (NHANES)

NHANES is an annual, nationally representative examination survey of the U.S. civilian, noninstitutionalized population. A stratified, multistage probability sampling scheme is used to select approximately 5,000 persons for personal interview and examination (22,23). The survey collects information on chronic disease prevalence (including undiagnosed conditions), risk factors, diet and nutritional status, immunization status, infectious disease prevalence, health insurance, and measures of environmental exposures.

Data from NHANES may be considered statistically unreliable for two reasons: sampling design and/or small sample size. Estimates for Healthy People 2010 objectives that are based on fewer than 30 sample events in the denominator or that have a RSE greater than 30 percent are also considered statistically unreliable and are suppressed. Other sources may use different criteria to assess the statistical reliability of NHANES data.

The NHANES sample size and design does not support producing estimates for certain racial and ethnic groups, including American Indians or Alaska Natives and Asians and Pacific Islanders. Healthy People 2010 tracking data for these populations are considered statistically unreliable and are suppressed. The survey is designed to provide an adequate sample size with adequate statistical coverage to produce estimates for Mexican Americans. While non-Mexican-American Hispanics are included in the sample, the (geographic) selection procedures do not provide for adequate statistical coverage of this ethnic group. Therefore, data for total Hispanics are shown in *Healthy People 2010* as DSU. More information on statistical reliability for data from this survey is available in a report by the Human Nutrition Information Service and the National Center for Health Statistics (12).

National Hospital Discharge Survey (NHDS)

NHDS is an annual survey that collects data from medical records to provide national estimates on hospital discharges from short-stay, noninstitutional hospitals and general and children's general hospitals regardless of length of stay (24,25). Annually, the national estimate is based on a sample of about 300,000 records. Key variables include length of stay, discharge status, International Classification of Diseases (ICD) diagnosis and procedure code, source of payment, hospital size, ownership, and location. Population-based rates are considered unreliable and are suppressed if the numerator is based on fewer than 30 records or if they have a RSE greater than 30 percent. Estimates based on 30–59 patient records are flagged to indicate they also have low reliability. Two Healthy People 2010 objectives based on the NHDS (objectives 5.10 and 16.5) use sample-based denominators to produce compound ratios. The determination of statistical reliability of these compound ratios involves consideration of additional criteria concerning the stability of the denominator. The additional criteria are specifically discussed in the updated operational definitions available online in DATA2010 (see section on DATA2010). Data on race and Hispanic origin are not available from some hospitals because the hospitals provide data from administrative data sets, which may not include race and ethnicity as required items. Hispanic origin was not reported for 75 percent of the NHDS records in 1992 (26). For all objectives tracked by this survey, data for American Indians or Alaska Natives and Asians or Pacific Islanders are statistically unreliable because of their small numbers in the survey. Data for Hispanics are suppressed because of high item nonresponse to the Hispanic origin variable.

National Household Survey on Drug Abuse (NHSDA)

NHSDA is an annual, multistage national probability sample survey of the civilian, noninstitutionalized population (ages 12 years and over) (27,28). In 1999, the sample size increased from about 25,000 to about 70,000 persons. The survey collects information on initiation, recency, and frequency of use of alcohol, tobacco (including smokeless

cigarettes and cigars), marijuana and other illicit drugs, prescription drug misuse, and treatment and prevention-related items. Estimated proportions are considered statistically unreliable and are suppressed if their RSE is greater than 17.5 percent. Estimated proportions are also considered statistically unreliable if $p < 0.0005$ or $p \geq 0.99995$.

National Survey of Family Growth (NSFG)

NSFG is a periodic survey based on a multistage probability design. It collects data on civilian, noninstitutionalized females 15–44 years of age by computer-assisted personal interview with a self-administered audio section for more sensitive topics (29–31). The sample size was 13,795 females in 1995. In 2002 civilian, noninstitutionalized males 15–44 years of age will be added to the sample. The content of this survey includes factors affecting birth and pregnancy, adoption, and maternal and infant health. Data are considered statistically unreliable when RSE is greater than 30 percent or the denominator is based on fewer than 50 sample cases. For all objectives tracked by this survey, data for American Indians or Alaska Natives and Asians or Pacific Islanders are statistically unreliable and are suppressed because the size of the sampled populations for these groups is too small.

1999 National Worksite Health Promotion Survey (NWHPS)

The 1999 NWHPS, a one-time telephone interview survey, defined a worksite as an organizational entity comprising 50 or more employees working at a particular geographic location. The sample was drawn randomly from the Dun & Bradstreet universe of employers (32). Approximately 1,500 worksites were surveyed in 1998–99. Key variables included employers' health risk and prevention programs and policies, corporate characteristics, and use of health plans for health promotion delivery. There were no specific guidelines employed to suppress data due to statistical unreliability, data quality, or confidentiality.

School Health Policies and Programs Study (SHPPS)

The periodic study—conducted every 6 years—consists of a census of all State education agencies; a national probability sample of public and private school districts; a national sample of public and private elementary, middle/junior high, and senior high schools; and a random sample of required health education and physical education classes. In 2000 all 51 State education agencies, 560 school districts, and 950 schools participated. Respondents included administrators, teachers, school nurses, counselors, food service staff, secretaries, and other school personnel (33). Through self-administered questionnaires and structured personal interviews, SHPPS collects data on the characteristics of school health programs, including health education, physical education, food service, health services, and school health policies. Data based on fewer than 30

schools in the denominator are considered statistically unreliable and are suppressed.

Youth Risk Behavior Surveillance System (YRBSS)

The national Youth Risk Behavior Survey (YRBS) is a part of the YRBSS and is a biennial, school-based survey administered to samples of students in grades 9–12. In 1999 about 15,300 students participated. The survey collects data on six categories of health risk behaviors: injury, tobacco use, alcohol and other drug use, sexual behavior, diet and nutrition, and physical activity (34,35). Data based on a denominator of fewer than 100 students are considered statistically unreliable and are suppressed. The national YRBS does not sample enough Asian or Pacific Islander or American Indian or Alaska Native adolescents from a single data year to present estimates for these race groups.

Population-Count Systems

Data from the systems discussed below are based on population counts of events and, as such, are not subject to sampling variability. However, data from these systems are subject to issues of data quality (for example, underreporting and misreporting) and random variation.

Most of the restrictions placed on showing data from these systems involve data based on a small number of cases for subnational geographic areas to address confidentiality concerns. These restrictions are generally not applicable to the Healthy People 2010 tracking data at the national level but are included in the discussion as guidance for those using these data for smaller geographic areas.

HIV/AIDS Surveillance System (HIV-AIDS)

The HIV-AIDS Surveillance System receives HIV and AIDS case and death reports from State and Territorial health departments on a voluntary basis. The data reports include information on the mode of exposure to HIV, case-definition category, and other clinical and demographic variables. Data from HIV-AIDS are not suppressed due to statistical unreliability; however, they are occasionally suppressed in order to protect case confidentiality (36). Data for selected counties or health districts and certain metropolitan statistical areas (MSAs) with populations between 100,000 and 500,000 are released in accordance with individual State's data release policies. For counties or health districts in selected States, only single-variable tables showing age, race/Hispanic origin, or sex can be released. For MSAs with populations between 100,000 and 500,000 in selected States, only two-variable tables showing selected variables can be released. For MSAs with populations less than 500,000 and where county or health district data are made available by State agreement, data for tabulations resulting in three or fewer cases are suppressed or noted as "less than or equal to three." For any State, the District of Columbia, or MSA with more than 500,000 persons, multiple-variable tables showing selected variables for adults and adolescents (13 years and over) are allowed (37).

National Notifiable Diseases Surveillance System (NNDSS)

NNDSS is an electronic surveillance system that compiles data submitted weekly by State and Territorial health departments on infectious diseases that have been designated nationally notifiable by CDC and the Council of State and Territorial Epidemiologists (CSTE). Although reporting of nationally notifiable infectious diseases to CDC by the States is voluntary, most States provide reports for these diseases as well as the internationally quarantinable diseases (cholera, plague, and yellow fever). Data from NNDSS are not suppressed due to statistical unreliability. However, demographic data (particularly race and Hispanic origin) are occasionally suppressed in order to protect case confidentiality. In accordance with CDC and CSTE data release policies, demographic breakdowns of data by county or any other geographic unit smaller than State are not allowed (38). In addition, if the total number of cases in a State in a year for a given condition is less than or equal to three, then race and Hispanic origin are suppressed for all cases in that State in that year.

National Profile of Local Health Departments (NPLHD)

NPLHD is a periodic census of approximately 2,800 local health departments (LHDs) including town/township, city, county, city-county, multicounty, and district health departments (39). Self-administered questionnaires are used to obtain information on the type of jurisdiction, services provided, staff size, community partnerships, managed care, and expenditures of each LHD. There is no sampling variability associated with NPLHD because the entire population of LHDs is queried. There are no formal criteria for classifying data as statistically unreliable. All data are presented regardless of item response rates or cell counts.

National Vital Statistics System-Nativity (NVSS-N) and National Vital Statistics System-Mortality (NVSS-M)

NVSS-N and NVSS-M contain data from birth and death certificates that are provided by the States, compiled and prepared by the National Center for Health Statistics, and presented in national statistical files (40). Data for Healthy People 2010 are based on resident births and deaths filed in the 50 States and the District of Columbia. Births or deaths of nonresidents of the United States are not included. Although the data are not subject to sampling error, counts may be affected by random variation. When data from this system are used for analytic purposes, such as the comparison of rates over time or for different areas, the number of events that actually occurred may be considered as one of a large series of possible events that could have arisen under the same circumstances. Rates, proportions, and simple ratios from NVSS are considered statistically unreliable and are suppressed if they are based on fewer than 20 events, which corresponds to an RSE of 23 percent or greater. For age-adjusted death rates, this criterion is applied to the sum of the age-specific deaths (41). Two

Healthy People 2010 objectives based on NVSS (objectives 5.6 and 5.7) use sample-based denominators from NHIS to produce compound ratios. The determination of statistical reliability of these compound ratios involves consideration of additional criteria concerning the stability of the denominator. The additional criteria are specifically discussed in the updated operational definitions available online in DATA2010 (see section on DATA2010).

Sexually Transmitted Diseases Surveillance System (STDSS)

STDSS obtains information on new sexually transmitted disease cases identified in the United States (42). Summary data reports, compiled by State and Territorial health departments from health care provider reports, are submitted to CDC on monthly, quarterly, and annual bases. Data from the STDSS are not suppressed due to statistical unreliability. Data are, however, occasionally suppressed in order to ensure case confidentiality. Data release policies (43) developed by the National Center for HIV, STD, and TB Prevention prohibit county-level tabulations of data unless prior consent is obtained from the State. In instances where permission is obtained, county-specific frequencies based on one to five cases are reported as “less than or equal to five” and rates are not reported. Demographic breakouts for State-level data can be shown, but counts of three or fewer are reported grouped or as “less than or equal to three.” National-level data are not subject to these restrictions.

United States Renal Data System (USRDS)

USRDS is a continuously updated database containing patient and facility records from the Centers for Medicare and Medicaid Services’ End Stage Renal Disease (ESRD) Program Management and Medical Information System and the Annual Facility Survey. The system provides national, State, and county estimates on ESRD, treatment modality, causes of death, patient survival, hospitalization, cost and cost effectiveness, and institutional providers of the ESRD treatment (44). Data are presented regardless of the value of their standard errors, which are presented with the data. Generally, data based on fewer than 10 persons or fewer than five providers are not shown due to confidentiality of patient- and provider-specific data.

Appendix

Race and Hispanic Origin

On October 30, 1997, the Office of Management and Budget (OMB) published *Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity* (45). The new standards, which revise OMB directive No. 15, adopted on May 12, 1977, reflect the increasing diversity of the Nation’s population (46). The U.S. Census Bureau used these standards in the 2000 decennial census. Other Federal programs will adopt the standards as soon as possible, but not later than January 1, 2003, for use in household surveys, administrative forms and

records, and other data collections. The major changes to the reporting of data on race and Hispanic origin under the new guidelines are 1) the instruction to “mark one or more” racial categories; 2) the division of the Asian or Pacific Islander category into two separate categories—“Asian” and “Native Hawaiian or Other Pacific Islander”; and 3) the renaming of the Black and Hispanic categories to “Black or African American” and “Hispanic or Latino,” respectively.

Healthy People 2010 adopted the new terminologies for the racial and ethnic categories and divided the Asian or Pacific Islander category into two categories as directed. However, because the OMB guidelines have generally not yet been implemented by most major health data systems, data for more than one race are not yet shown in the Healthy People 2010 tracking data. When they become available, data for more than one race will be included in the tracking data for most of the major Healthy People 2010 data systems. In addition, many data systems currently do not produce data separately for the Asian and Native Hawaiian or Other Pacific Islander populations. In these cases data will continue to be shown for the combined category Asian or Pacific Islander where available. Until data are collected separately for the Asian and the Native Hawaiian or Other Pacific Islander populations by these data systems, they are shown as not collected (DNC).

As data systems adopt the new collection standards, there may be changes in the ability of the data systems to report estimates for the newly defined groups. This will be monitored closely over the decade.

Resources

Healthy People 2010 and Tracking Healthy People 2010

Healthy People 2010 (3) provides a comprehensive overview of the Healthy People 2010 initiative. Published as a two-volume set in November 2000, these reference books contain detailed descriptions of the 467 objectives. Volume 1 discusses the history of *Healthy People 2010*, the development of the Leading Health Indicators, and contains information on objectives in focus areas 1–14. Volume 2 contains information on objectives in focus areas 15–28. For each objective, the target and data source are shown and all available data for the baseline year are provided.

Another available resource is *Tracking Healthy People 2010* (4), a technical guidebook on the statistical measures used for *Healthy People 2010*. This book contains operational definitions for each of the objective and subobjective measures. It also contains profiles of the 23 major data sources used, including information on survey design and content, mode of administration, periodicity, and availability of geographic estimates. A discussion of general data issues affecting many objectives (such as age adjustment, morbidity and mortality classification, variability of estimates, target setting methods, strategies for assessing progress, etc.) is included. Contact information for the agencies responsible for the objectives is also provided. The guidebook was developed to assist Healthy People 2010

programs at the national, State, and local levels, but may also be of interest to those who wish to better understand the analytic framework of *Healthy People 2010*.

Healthy People 2010 and *Tracking Healthy People 2010* are available on the Internet at www.health.gov/healthypeople/publications. Hard copies may be ordered from the U.S. Government Printing Office by calling (202) 512-1800 or by submitting an electronic request at bookstore.gpo.gov.

DATA2010

The Healthy People 2010 database, called “DATA2010,” contains baseline and tracking data for each of the Leading Health Indicators and the Healthy People 2010 objective and subobjective measures. Users can generate tables that display data for a single objective, an entire focus area, a specific population such as a race, Hispanic origin, or gender group, or all objectives tracked by a particular data system. Data for the ten Leading Health Indicators designated by DHHS can also be shown separately (3). These tables include the data source, 2010 target, and data for each group in the population template for the baseline year and all subsequent years (as available). While national data are available for all objectives, State data are currently available for a subset of the Healthy People 2010 objectives. DATA2010 also includes updates and revisions to the operational definitions published in *Tracking Healthy People 2010*. DATA2010 can be accessed online through the CDC WONDER system at wonder.cdc.gov/data2010 or through the NCHS Web site at www.cdc.gov/nchs/hphome.htm.

Data updates and revisions, including changes to baselines and targets, are posted to DATA2010 on a quarterly basis. A cumulative report of baseline data changes made since the publication of *Healthy People 2010* in November 2000—including reasons for changes—is available when creating tables in DATA2010 or can be accessed directly at wonder.cdc.gov/data2010/bslrevis.htm.

References

1. Department of Health and Human Services. Promoting health/preventing disease: Objectives for the nation. Washington: Public Health Service. 1980.
2. Department of Health and Human Services. Healthy People 2000: National health promotion and disease prevention objectives. Washington: Public Health Service. 1991.
3. Department of Health and Human Services. Healthy People 2010. 2nd ed. With understanding and improving health and objectives for improving health. 2 vols. Washington: U.S. Government Printing Office. November 2000.
4. Department of Health and Human Services. Tracking Healthy People 2010. Washington: U.S. Government Printing Office. November 2000.
5. Curtin LR, Klein RJ. Direct standardization (age-adjusted death rates). Healthy People 2000 Statistical Notes, no. 6-revised. Hyattsville, Maryland: National Center for Health Statistics. 1995.
6. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People 2010 statistical notes, no 20. Hyattsville, Maryland: National Center for Health Statistics. 2001.
7. Nelson DE, Holtzman D, Walker M, et al. Objectives and design of the Behavioral Risk Factor Surveillance System. Proceedings of the Survey Research Methods section of the American Statistical Association Joint Statistical Meetings. Alexandria, Virginia: American Statistical Association. 1998.
8. Powell-Griner E. Uses and limitations of the Behavioral Risk Factor Surveillance System data. Proceedings of the Survey Research Methods section of the American Statistical Association Joint Statistical Meetings. Alexandria, Virginia: American Statistical Association. 1998.
9. Tippet KS, Cypel YS, eds. Design and Operation: The Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey, 1994–96. U.S. Department of Agriculture, Nationwide Food Surveys Report No. 96-1. 1998.
10. U.S. Department of Agriculture, Agricultural Research Service. Continuing Survey of Food Intakes by Individuals, 1994–96 and 1998. CD-ROM. 2000.
11. Life Sciences Research Office, Federation of American Societies for Experimental Biology. Prepared for the Interagency Board for Nutrition Monitoring and Related Research. Third report on nutrition monitoring in the United States. Vol. I. Washington: U.S. Government Printing Office. 1995.
12. Human Nutrition Information Service/National Center for Health Statistics Analytic Working Group. Joint policy on variance estimation and statistical reporting standards on NHANES III and CSFII reports: HNIS/NCHS analytic working group recommendations. Hyattsville, Maryland: National Center for Health Statistics. 1993.
13. Cohen JW, Monheit AC, Beauregard KM, et al. The Medical Expenditure Panel Survey: A national health information resource. *Inquiry* 33:373–89. 1997.
14. Bachman JG, Johnston LD, O’Malley PM. Monitoring the future: Questionnaire responses from the Nation’s high school seniors 2000. Ann Arbor, Michigan: Institute for Social Research. 2001.
15. Johnston LD, O’Malley PM, Bachman JG. National survey results on drug use from the Monitoring the Future study, 1975–2000. Volume I: Secondary school students, and Volume II: College students and adults through age 40. Bethesda, Maryland: National Institute on Drug Abuse. 2001.
16. Woodwell DA. National Ambulatory Medical Care Survey: 1997 summary. Advance data from vital and health statistics; no 305. Hyattsville, Maryland: National Center for Health Statistics. 1999.
17. McCaig LF, McLemore T. Plan and operation of the National Hospital Ambulatory Medical Care Survey. National Center for Health Statistics. *Vital Health Stat* 1(34). 1994.
18. Hubble D. The National Crime Victimization Survey redesign: New questionnaire and procedures development and phase-in methodology. Orlando, Florida: Annual meeting of the American Statistical Association. 1995.
19. Kinderman C, Lynch J, Cantor D. Effects of the redesign on victimization estimates. Washington: Bureau of Justice Statistics. 1997.
20. Hendershot G, Adams P, Marano M, et al. Current estimates from the National Health Interview Survey, 1996. National Center for Health Statistics. *Vital Health Stat* 10(200). 1999.

21. Massey JT, Moore TF, Parsons VL, et al. Design and estimation for the National Health Interview Survey, 1985–94. National Center for Health Statistics. *Vital Health Stat* 2(110). 1989.
22. National Center for Health Statistics. Plan and operation of the third National Health and Nutrition Examination Survey, 1988–94. National Center for Health Statistics. *Vital Health Stat* 1(32). 1994.
23. Ezzati TM, Massey JT, Waksberg J, et al. Sample design: Third National Health and Nutrition Examination Survey. National Center for Health Statistics. *Vital Health Stat* 2(113). 1992.
24. Graves EJ, Owings MF. 1996 Summary: National Hospital Discharge Survey. Advance data from vital and health statistics, no 301. Hyattsville, Maryland: National Center for Health Statistics. 1998.
25. Dennison C, Pokras R. Design and operation of the National Hospital Discharge Survey: 1988 redesign. National Center for Health Statistics. *Vital Health Stat* 1(39). 2000.
26. Kozak LJ. Underreporting of race in the National Hospital Discharge Survey. Advance data from vital and health statistics, no 265. Hyattsville, Maryland: National Center for Health Statistics. 1995.
27. Substance Abuse and Mental Health Services Administration. Development and implementation of a new data collection instrument for the 1994 National Household Survey on Drug Abuse. Washington: U.S. Government Printing Office. 1996.
28. Gfroerer J. An overview of the National Household Survey on Drug Abuse and related methodological research. Proceedings of the Survey Research section of the American Statistical Association Joint Statistical Meetings. Boston, Massachusetts. 1992.
29. Kelly JE, Mosher WD, Duffer AP, et al. Plan and operation of the 1995 National Survey of Family Growth. National Center for Health Statistics. *Vital Health Stat* 1(36). 1997.
30. Potter FJ, Iannacchione VG, Mosher WD, et al. Sampling weights, imputation, and variance estimation in the 1995 National Survey of Family Growth. National Center for Health Statistics. *Vital Health Stat* 2(124). 1998.
31. Abma J, Chandra A, Mosher W, Peterson L, Piccinino L. Fertility, family planning, and women's health: New data for the 1995 National Survey of Family Growth. National Center for Health Statistics. *Vital Health Stat* 23(19). 1997.
32. Association for Worksite Health Promotion. 1999 National Worksite Health Promotion Survey. Northbrook, Illinois: The Association. 1999.
33. Smith T, Brener N, Kann L, Kinchen S, et al. Methodology for the School Health Policies and Programs Study 2000. *J Sch Health* 71(7):260–5. 2001.
34. Centers for Disease Control and Prevention. Youth risk behavior surveillance, United States. *MMWR* 49(SS-5):1–96. 2000.
35. Kann L, Kolbe LJ, Collins JL, eds. Measuring the health behavior of adolescents: The Youth Risk Behavior Surveillance System and recent reports on high-risk adolescents. *Public Health Rep* 108(Supp 1). 1993.
36. Centers for Disease Control and Prevention. Guidelines for national human immunodeficiency virus case surveillance, including monitoring for human immunodeficiency virus infection and acquired immunodeficiency syndrome. *MMWR* 48(RR13). 1999.
37. National Center for HIV, STD, and TB Prevention. AIDS public information data set manual. Atlanta, Georgia: Centers for Disease Control and Prevention. 1999.
38. Epidemiology Program Office. Data release guidelines for the National Notifiable Diseases Surveillance System. Atlanta, Georgia: Centers for Disease Control and Prevention. 2000.
39. National Association of County and City Health Officials. 1992–93 National Profile of Local Health Departments: National surveillance series. Atlanta, Georgia: Centers for Disease Control and Prevention. 1995.
40. National Center for Health Statistics. Vital statistics of the United States, vol I, natality. Technical Appendix. Hyattsville, Maryland: National Center for Health Statistics. 1999. Available at: www.cdc.gov/nchs.
41. National Center for Health Statistics. Vital statistics of the United States, vol II, mortality. Technical Appendix. Hyattsville, Maryland: National Center for Health Statistics. 1995. Available at: www.cdc.gov/nchs.
42. National Center for HIV, STD, and TB Prevention. Sexually transmitted disease prevention surveillance. Atlanta, Georgia: Centers for Disease Control and Prevention. 1999.
43. National Center for HIV, STD, and TB Prevention. STD surveillance data use and release policy. Atlanta, Georgia: Centers for Disease Control and Prevention. 1999.
44. United States Renal Data System. 1999 annual data report. Bethesda, Maryland: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases. Chapter XIII. 1999.
45. Office of Management and Budget. Standards for maintaining, collecting, and presenting Federal data on race and ethnicity. *Federal Register*, 62 FR 58781–58790. 1997.
46. Office of Management and Budget. Directive No. 15. Race and ethnic standards for Federal statistics and administrative reporting. *Statistical Policy Handbook*. Washington: Office of Management and Budget. 1978.

Table 1. Summary of criteria for data suppression for the *Healthy People 2010* major data systems

Data system	Number of HP2010 objectives tracked	Criteria for DSU ¹	Population groups always displayed as DSU in HP2010 ²
<i>Sample surveys</i>			
Behavioral Risk Factor Surveillance System (BRFSS)	9	Denominator < 50	(None)
Continuing Survey of Food Intake by Individuals (CSFII)	6	Means: Sample size < 30•VIF or CV > 30% Estimates 25–75%: Sample size < 30•VIF or CV > 30% Estimates < 25% or > 75%: Smaller of np or n(1 – p) < 8•VIF	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander
Medical Expenditure Panel Survey (MEPS)	6	Denominator < 70; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander
Monitoring the Future Study (MTF)	5	No formal criteria	American Indian or Alaska Native Asian
National Ambulatory Medical Care Survey (NAMCS)	7	Numerator < 30; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander Hispanic or Latino
National Crime Victimization Survey (NCVS)	6	Numerator < 10	(None)
National Health Interview Survey (NHIS)	67	Denominator < 50; or RSE > 30%	(None)
National Health and Nutrition Examination Survey (NHANES)	35	Denominator < 30; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander Hispanic or Latino
National Hospital Ambulatory Medical Care Survey (NHAMCS)	6	Numerator < 30; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander Hispanic or Latino
National Hospital Discharge Survey (NHDS)	11	Rates: Numerator < 30 records; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander Hispanic or Latino
National Household Survey on Drug Abuse (NHSDA)	8	RSE > 17.5%	(None)
National Survey of Family Growth (NSFG)	14	Denominator < 50; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander
1999 National Worksite Health Promotion Survey (NWHPS)	5	No formal criteria	(Data are not population based)
School Health Policies and Programs Study (SHPPS)	8	Denominator < 30 schools	(Data are not population based)
Youth Risk Behavior Surveillance System (YRBSS)	11	Denominator < 100 events	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander

See footnotes at end of table.

Table 1. Summary of criteria for data suppression for the *Healthy People 2010* major data systems—Con.

Data system	Number of HP2010 objectives tracked	Criteria for DSU ¹	Population groups always displayed as DSU in HP2010 ²
<i>Population-count systems</i>			
HIV/AIDS Surveillance System (HIV-AIDS)	10	< 4 cases	(None)
National Notifiable Diseases Surveillance System (NNDSS)	5	Race and Hispanic origin suppressed if < 4 cases	(None)
National Profile of Local Health Departments (NPLHD)	8	No formal criteria	(Data are not population based)
National Vital Statistics System-Mortality (NVSS-M)	32	< 20 deaths	(None)
National Vital Statistics System-Natality (NVSS-N)	8	< 20 births	(None)
STD Surveillance System (STDSS)	7	County: < 6 cases State: < 4 cases National: No restrictions	(None)
United States Renal Data System (USRD)	7	< 10 persons; or < 5 providers	(None)

¹See text for a more detailed discussion.

²Data system does not produce reliable estimates for these groups.

List of Abbreviations:

CV: Coefficient of variation.

DSU: Data statistically unreliable.

VIF: Variance inflation factor.

RSE: Relative standard error.

Table 2. Sponsoring agency and Internet address for *Healthy People 2010* major data systems

Data system	Sponsoring agency	Internet address
<i>Sample surveys</i>		
Behavioral Risk Factor Surveillance System (BRFSS)	HHS, CDC, NCCDPHP	www.cdc.gov/nccdphp/brfss
Continuing Survey of Food Intakes by Individuals (CSFII)	USDA, ARS	www.barc.usda.gov/bhnrc/foodsurvey/home.htm
Medical Expenditure Panel Survey (MEPS)	HHS, AHRQ	
Monitoring the Future Study (MTF)	HHS, CDC, NCHS	www.meps.ahrq.gov/survey.htm#target2
National Ambulatory Medical Care Survey (NAMCS)	HHS, NIH, NIDA	www.isr.umich.edu/src/mtf/index.html
National Crime Victimization Survey (NCVS)	HHS, CDC, NCHS	www.cdc.gov/nchs/about/major/ahcd/ahcd1.htm
National Health Interview Survey (NHIS)	DOJ, BJS	www.ojp.usdoj.gov/bjs/cvict.htm#ncvs
National Health and Nutrition Examination Survey (NHANES)	HHS, CDC, NCHS	www.cdc.gov/nchs/nhis.htm
National Hospital Ambulatory Medical Care Survey (NHAMCS)	HHS, CDC, NCHS	www.cdc.gov/nchs/about/major/ahcd/ahcd1.htm
National Hospital Discharge Survey (NHDS)	HHS, CDC, NCHS	www.cdc.gov/nchs/about/major/hdasd/nhds.htm
National Household Survey on Drug Abuse (NHSDA)	HHS, SAMHSA	www.samhsa.gov/oas/nhsda.htm
National Survey of Family Growth (NSFG)	HHS, CDC, NCHS	
	HHS, CDC, NICHD	
	HHS, CDC, NCHSTP	www.cdc.gov/nchs/nsfg.htm
	OPA	
1999 National Worksite Health Promotion Survey (NWHPS)	HHS, ODPHP	
	AWHP	
	William M. Mercer, Inc.	www.odphp.osophs.dhhs.gov
School Health Policies and Programs Study (SHPPS)	HHS, CDC, NCCDPHP	www.cdc.gov/nccdphp/dash/shpps/index.htm
Youth Risk Behavior Surveillance System (YRBSS)	HHS, CDC, NCCDPHP	www.cdc.gov/yrbs

See footnotes at end of table.

Table 2. Sponsoring agency and Internet address for *Healthy People 2010* major data systems—Con.

Data system	Sponsoring agency	Internet address
<i>Population-count systems</i>		
HIV/AIDS Surveillance System (HIV-AIDS)	HHS, CDC, NCHSTP	www.cdc.gov/nchstp/hiv_aids/surveillance.htm
National Notifiable Diseases Surveillance System (NNDSS)	HHS, CDC, EPO	www.cdc.gov/epo/dphsi/phs.htm
National Profile of Local Health Departments (NPLHD)	NACCHO	www.naccho.org
National Vital Statistics System-Mortality (NVSS-M)	HHS, CDC, NCHS	www.cdc.gov/nchs/about/major/dvs/mortdata.htm
National Vital Statistics System-Natality (NVSS-N)	HHS, CDC, NCHS	www.cdc.gov/nchs/births.htm
STD Surveillance System (STDSS)	HHS, CDC, NCHSTP	www.cdc.gov/nchstp/dstd/dstdp.html
United States Renal Data System (USRDS)	HHS, NIH, NIDDK HHS, CMS	www.usrds.org

List of abbreviations:

AHRQ: Agency for Healthcare Research and Quality
 ARS: Agricultural Research Service
 AWHP: Association for Worksite Health Promotion
 BJS: Bureau of Justice Statistics
 CDC: Centers for Disease Control and Prevention
 CMS: Centers for Medicare and Medicaid Services (formerly HCFA)
 DOJ: U.S. Department of Justice
 EPO: Epidemiology Program Office
 HCFA: Health Care Finance Administration (now CMS)
 HHS: U.S. Department of Health and Human Services
 NACCHO: National Association of County and City Health Officials
 NCCDPHP: National Center for Chronic Disease Prevention and Health Promotion
 NCHS: National Center for Health Statistics
 NIDA: National Institute on Drug Abuse
 NIDDK: National Institute of Diabetes and Digestive and Kidney Diseases
 NIH: National Institutes of Health
 NCHSTP: National Center for HIV, STD, and TB Prevention
 ODPHP: Office of Disease Prevention and Health Promotion
 OPA: Office of Population Affairs
 SAMHSA: Substance Abuse and Mental Health Services Administration
 USDA: U.S. Department of Agriculture

Recent issues of *Healthy People* Statistical Notes

Number	Title	Date of Issue
20	Age Adjustment Using the 2000 Projected U.S. Population	January 2001
21	Summary Measures of Population Health: Methods for Calculating Healthy Life Expectancy	August 2001
22	Summary Measures of Population Health: Addressing the First Goal of <i>Healthy People 2010</i> , Improving Health Expectancy	September 2001
23	Trends in Racial and Ethnic-Specific Rates for the Health Status Indicators: United States, 1990–98	January 2002

NOTES: Issues 1–19 and 23 address *Healthy People 2000*. Issues 20–22 and 24 address *Healthy People 2010*. Issues 6–24 are available on the Internet at www.cdc.gov/nchs.

**DEPARTMENT OF
HEALTH & HUMAN SERVICES**

**Centers for Disease Control and Prevention
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782-2003**

FIRST CLASS MAIL
POSTAGE & FEES PAID
CDC/NCHS
PERMIT NO. G-284

**OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, \$300**

**To receive this publication regularly, contact
the National Center for Health Statistics by
calling 301-458-4636
E-mail: nchsquery@cdc.gov
Internet: www.cdc.gov/nchs**

Suggested citation

Klein RJ, Proctor SE, Boudreault MA, Turczyn KM.
Healthy People 2010 criteria for data suppression.
Statistical Notes, no 24. Hyattsville, Maryland:
National Center for Health Statistics. June 2002.