Healthy People 2010:

General Data Issues

1. Introduction

In November 2000, the Healthy People 2010 objectives were published in two volumes [1]. Concurrently, a third volume, the statistical compendium *Tracking Healthy People* was also released [2]. The purpose of *Tracking Healthy People* was to provide technical detail for statistical and data-related issues that cut across many Healthy People 2010 Focus Areas and objectives.

This document provides the final update of *Tracking Healthy People, Part A: General Data Issues,* previously updated in 2006. This information complements the final evaluation of the Healthy People 2010 objectives, *Healthy People 2010 Final Review,* released in October 2011 [3].

Tracking Healthy People contained three major parts:

- Part A—General Data Issues—discussion of issues that affect many Healthy People 2010 objectives
- Part B—Operational Definitions—technical specifications for each Healthy People 2010 objective
- Part C—Major Data Sources for Healthy People 2010—detailed information on Healthy People 2010 data sources.

Tracking Healthy People was updated and released on the Healthy People 2010 online data tracking system, DATA2010, accessible from http://wonder.cdc.gov/data2010/, at the time of the publication of the *Healthy People 2010 Midcourse Review* in 2006 [4].

Updates for the technical specifications for specific objectives and major data sources sections of *Tracking Healthy People* are available from http://wonder.cdc.gov/data2010/technicl.htm.

The following topics are included in this document (topics added since the original publication of *Tracking Healthy People, Part A: General Data Issues*, are shown with an *):

- Measuring quality and years of healthy life*
- Measuring health disparities*
- Target setting and assessing progress for measurable objectives
- Population estimates
- Population template
- Age adjustment
- Mortality and morbidity classification
- National data
- State and local data
- Variability of estimates and data suppression
- Healthy People 2010 database

References

1. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.
2. U.S. Department of Health and Human Services, Public Health Service. Tracking Healthy People 2010. Washington, D.C.: Government Printing Office. November 2000.
3. National Center for Health Statistics. Healthy People 2010 Final Review. Hyattsville, MD. 2011.
4. U.S. Department of Health and Human Services. Healthy People 2010 Midcourse Review. Washington, D.C.: U.S. Government Printing Office. 2006.

2. Measuring Quality and Years of Healthy Life

Introduction

The first overarching goal of Healthy People 2010 is to help individuals of all ages increase life expectancy and improve their quality of life [1–3].

Life expectancy is the average number of years people born in a given year are expected to live based on a set of age-specific death rates. At the beginning of the 20th century, life expectancy at birth was 47.3 years. Life expectancy has increased dramatically over the past 100 years; in 2007 life expectancy at birth was 77.9 years [4].

The gains in life expectancy since the early 1900s are largely attributable to the control of infectious diseases through improved sanitation, vaccines, and antimicrobials; improved nutrition; and advances in medical research and treatment. However, longevity is no longer a sufficient measure of the health of a population. Over the last century, chronic diseases have replaced infectious diseases as major causes of death, resulting in an increase in the number of persons with functional limitations associated with chronic illness. Preventing disabling conditions, improving function, relieving physical pain and emotional distress, and maximizing health across the life span have become issues that are as important to address as increasing life expectancy [5].

Measuring Life Expectancy

Life expectancy is measured by constructing a life table. There are two types of life tables: the cohort (or generation) life table and the period (or current) life table. The cohort life table presents the mortality experience of a particular birth cohort (for example, all persons born in the year 1900) from birth throughout their lives. The cohort life table is based on age-specific death rates observed throughout the lifetime of the cohort members and thus reflects the mortality experience of an actual population from birth until the final group member has died [4].

Unlike the cohort table, the period table does not represent the mortality experience of an actual birth cohort. Rather, the period table presents what would happen to a hypothetical cohort if it experienced throughout its entire life, the mortality conditions of a particular period in time. Thus, for example, a period life table for 2007 assumes a hypothetical cohort subject throughout its lifetime to the age-specific death rates prevailing for the current population in 2007 [4]. The period table is used to construct the life expectancies tracked in *Healthy People 2010* [2,3]. The methodology for constructing period life tables for the United States has been published elsewhere [6].

Measuring Quality and Years of Healthy Life

Due to the multi-dimensional nature of health, assessing quality and years of healthy life is more complex than measuring life expectancy, and the field is still evolving. Various measures are used nationally and internationally to measure healthy life. These fall into three general categories [7]:

- Self-assessments of overall health status by individuals or their proxies
- Composite measures that include multiple dimensions of health. Scores on the various dimensions are combined into a single measure using a predetermined algorithm
- Measures that combine mortality and morbidity (where the morbidity measure can be either of the types described above or a measure of a single dimension of health); these measures use years as the metric to quantify healthy life.

Healthy People 2010 mentioned several measures of population health: self-perceived health status, healthy days, and Years of Healthy Life (YHL), the measure used in Healthy People 2000 [1–3,8]. In 1998, the National Center for Health Statistics convened a workshop to select measures that best capture the complexity of assessing years of healthy life within the context of Healthy People 2010 [9]. Three summary measures that combine mortality with different measures of morbidity were selected to track progress towards Goal 1 of Healthy People 2010. The measures are:

1. Expected years in good or better health
2. Expected years free of activity limitation
3. Expected years free of chronic disease.

The first two summary measures evolved from the YHL measure. YHL combines information about mortality, self-rated health, and activity limitation into a single measure. The new measures separate the self-rated health component from the limitation of activities component to better track change over time. Expected years of life free of chronic diseases was added to provide an additional aspect of population health.

Healthy Life Measures

The three healthy life measures used in Healthy People 2010 are defined as follows [2,3]:

Expected years in good or better health is the average number of years a person can expect to live in good or better health. This measure assesses healthy life using a single global assessment question which asks a person to rate his or her health as “excellent,” “very good,” “good,” “fair,” or “poor”.

Expected years free of activity limitation is the average number of years a person can expect to live free from a limitation in activities, a need for assistance in personal or routine care needs, or a need to use special equipment.

Expected years free of chronic disease is the average number of years a person can expect to live without developing one or more of the following conditions: heart disease, cancer, diabetes, hypertension, kidney disease, arthritis, or asthma.

Computing Healthy Life Measures

Healthy life expectancies are calculated using a period life table methodology as described above [6]. Age-specific mortality rates are combined with age-specific prevalence rates to produce an estimate of overall healthy life expectancy. The methodology has been published elsewhere [9].

Healthy life expectancies can be compared across populations, as well as over time, as long as the age-specific prevalence rates are reliable across all age categories. Often, several years of data must be combined to produce reliable rates.

Data Sources for Healthy Life Measures

Healthy life calculations use mortality data from the National Vital Statistics System and health data from the National Health Interview Survey (NHIS). (See *Tracking Healthy People, Part C*, for descriptions of these systems.) Although the NHIS provides detailed information on health and health behaviors, the institutionalized population is excluded from the NHIS sample. Since the institutionalized population is more likely to report poor health, measures may underestimate the effect of the health component on healthy life expectancies and, therefore, may overestimate healthy life expectancy.

Data items used for Healthy Life measures

Self-rated health status is measured by the single question in which the respondent is asked to rate his or her health as “excellent,” “very good,” “good,” “fair,” or “poor.” This self-assessed health rating has been validated to be a useful indicator of a person’s actual health for a variety of populations and thus permits broad comparisons across different conditions and populations [10]. In addition to the NHIS, the National Health and Nutrition Examination Survey, the Behavioral Risk Factor Surveillance System, and other surveys include the measure.

Activity limitation is measured using questions about personal care needs, limitations of activities, and use of special equipment. Adults are asked whether they need assistance with personal care needs, such as eating, bathing, dressing, or getting around inside the home; if they need assistance with routine care needs, such as household chores; if they have mental or physical problems that prevent them from working at a job; or if they have health problems that require the use of special equipment, such as a cane, wheelchair, or special telephone. Persons responding “yes” to any of these questions are classified as having an activity limitation. Children are considered limited in activity if the proxy adult respondent answers "yes" to any of the limitation, special services, or special equipment questions.

Chronic disease status is measured by a question which asks if a doctor has ever diagnosed the respondent with a given disease. The list of selected chronic diseases includes hypertension, stroke, coronary heart disease, arthritis, kidney disease, diabetes, and cancer. All selected chronic diseases have related objectives within Healthy People 2010. Respondents who answer “yes” to any of the selected diagnoses are classified as having a chronic disease. The NHIS does not request information on the severity of the disease, even though relative risk of mortality varies with the severity and type of chronic disease.

Future Directions

Since the measurement of healthy life is an evolving field, it is important to continue to develop better measures of healthy life and to improve the data on which these measures are based.

Refining Measurement of Quality and Years of Healthy Life

The three summary measures selected for use in *Healthy People 2010* have provided trend data for the final decade review [3]. However, measures incorporating other aspects of health that have not been included in summary measures of health, such as healthy behaviors and mental health, are needed to provide different ways of summarizing this important construct. Work is continuing in this area.

Expanding Data Collection

Better data are necessary to support both the current measures and any newly developed measures. Limited socio-demographic data are a particular problem. Currently, the black and white populations are the only population groups with sufficiently reliable data in most national data sources to produce years of healthy life estimates across the decade. Life tables for the Hispanic population have become available only since 2006 [4]. In addition, social determinants of health, such as education, income, or occupation, are sometimes lacking from national health data sources. For the Healthy People objectives, understanding the status of subgroups of the population is important for public health policy.

Another challenge is collection of data on the institutionalized populations such as those in prisons and nursing homes. Household-based surveys, the source of much of the data used to measure the health components of healthy life expectancy, do not collect information on these populations or on the homeless. Many of these individuals are likely to experience poor health, and estimates of healthy life that do not include these populations are biased.

References

1. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.
2. U.S. Department of Health and Human Services. Healthy People 2010 Midcourse Review. Washington, D.C.: U.S. Government Printing Office. 2006.
3. National Center for Health Statistics. Healthy People 2010 Final Review. Hyattsville, MD. 2011.
4. Arias E. United States life tables, 2007. Natl Vital Stat Rep 59(9). Hyattsville, MD: National Center for Health Statistics. 2011. Available from: http://www.cdc.gov/nchs/data/nvsr/nvsr59/nvsr59_09.pdf.
5. Molla MT, Wagener DK, Madans JH. Summary measures of population health: Methods for calculating healthy life expectancy. Healthy People Statistical Notes, no. 21. Hyattsville, MD: National Center for Health Statistics. August 2001. Available from: www.cdc.gov/nchs/data/statnt/statnt21.pdf.
6. Anderson RN. Method for constructing complete annual U.S. life tables. National Center for Health Statistics. Vital Health Stat 2(129). 1999. Available from: www.cdc.gov/nchs/data/series/sr_02/sr02_129.pdf.
7. Murray CJL, Salomon JA, Mathers CD, Lopez AD (eds). Summary Measures of Population Health: Concepts, Ethics, Measurement, and Applications. Geneva, Switzerland: World Health Organization. 2002.
8. Erickson P, Wilson R, Shannon, I. Years of healthy life. Healthy People Statistical Notes, no. 7. Hyattsville, MD: National Center for Health Statistics. 1995. Available from: www.cdc.gov/nchs/data/statnt/statnt07.pdf.
9. Molla MT, Madans JH, Wagener DK, Crimmins EM. Summary measures of population health: Report of findings on methodologic and data issues. National Center for Health Statistics. Hyattsville, MD. 2003. Available from: www.cdc.gov/nchs/data/misc/pophealth.pdf.
10. Idler EL, Benyamini Y. Self-rated health and mortality: A Review of twenty-seven community studies. J Health Soc Behav 38: 21–37. 1997.

3. Measuring Health Disparities

The second overarching goal of Healthy People 2010 is to eliminate disparities among segments of the population, including differences that occur by sex, race or ethnicity, education or income, disability, geographic location or sexual orientation [1–3]. The specific population groups associated with each of these characteristics are defined in Section 6: Population Template. The goal of eliminating health disparities applies to 498 population-based Healthy People 2010 objectives; that is, those objectives that measure health aspects of the population. The goal is not applicable to objectives that measure schools, worksites, or states or to objectives that are tracked using counts of events rather than rates or percents.

In *Healthy People 2010*, disparity is defined as the quantity that separates a group from a reference point on a particular measure of health that is expressed in terms of a rate, proportion, mean, or some other quantitative measure [2–4]. In order to measure disparity between groups, four critical analytic issues need to be addressed:

- What reference point should be used to measure disparity?
- Should disparity be measured in absolute or relative terms?
- Should disparity be measured in terms of favorable or adverse events?
- How can disparity be summarized across multiple subgroups?

A detailed discussion of these conceptual issues and their implications for measuring disparity in *Healthy People 2010* has been published [2–4]. A brief synopsis of the issues and the approaches adopted for *Healthy People 2010* is provided below.

Choice of reference point

Disparity can be measured in relation to various reference points, including the year 2010 target, the total population rate, an average of group rates, or the rate for a specific group such as the largest group, or the group with the most favorable or "best" rate. There are advantages and disadvantages associated with each of these reference points.

The "best" group rate has been chosen as the reference point for measuring disparities in *Healthy People 2010* because it emphasizes the potential for improvement among the comparison groups. Implicit in the use of the "best" group rate as the reference point is the idea that the best group rate is theoretically achievable by the other groups [2–4].

The best group rate is used as the reference point in analyses of characteristics with two groups (e.g., sex, disability, geographic location) and characteristics with three or more groups (e.g., race and ethnicity, education, or income). The disparity for a particular characteristic is measured by comparing the best group rate and the rates for the other groups at a given point in time. It is important to remember that, for a particular objective, the best rate will differ from one characteristic to the next. In addition, for a particular characteristic, the group with the best rate may change over time.

In some instances, the group with the most favorable rate is not suitable for selection as the reference point because it does not meet an additional criterion for statistical reliability. In *Healthy People 2010* disparity analyses, a group can only be selected as the reference point if its rate has a relative standard error of less than 10 percent [2,3]. This criterion is more stringent than the 30 percent relative standard error used by many Healthy People 2010 data sources for data suppression [5]. This criterion prevents the measurement of disparity from a highly variable reference point. If the group with the most favorable rate does not meet the criterion for selection as the reference point, then the group with the next most favorable rate that meets this criterion is selected. If no more than one group meets this criterion, measures of disparity are not computed for that characteristic.

Measuring disparity on an absolute or relative scale

Differences between the best group rate and the other group rates can be measured in absolute or relative terms. Absolute measures, such as the simple difference, describe the arithmetic difference between group rates and are expressed in the same units of measurement as the group rates. Relative measures, such as the percent difference, describe the arithmetic difference between group rates relative to a reference rate – that is, as a percentage of the reference rate. Relative measures are unitless. Absolute measures are valuable in assessing public health impact and can be easier to interpret than relative measures. However, relative measures make it possible to compare disparities across objectives that are measured on different scales.

In general, absolute and relative measures yield concordant conclusions about disparity at a *point in time*. However, in some instances absolute and relative measures of disparity may lead to different conclusions about changes *over time*. For example, when the best group rate is declining, a reduction in the absolute difference from the best group rate can occur without a reduction in the relative difference. Relative statistics are more appropriate for trend analyses because they adjust for changes in the level of the reference point over time and adjust for differences in the level of the reference point when comparisons are made across objectives.

Measuring disparity in terms of favorable or adverse events

Some Healthy People 2010 objectives are expressed in terms of favorable events or conditions that are to be increased while others are expressed in terms of adverse events or conditions that are to be reduced. For example, objective 1-1 is expressed in terms of favorable events: *increase the proportion of persons with health insurance*. Conversely, objective 3-1 is expressed in terms of adverse events: *reduce the overall cancer death rate*.

The magnitude of an absolute measure of disparity at a particular point in time does not depend on whether an indicator is expressed in terms of adverse or favorable events. However, the magnitude of a relative measure of disparity does depend on the way that the objective is expressed. In addition, conclusions about *changes in disparity over time* depend on whether the objective is expressed in terms of favorable or adverse events. A more detailed explanation, including numerical examples, has been published [4,6].

For these reasons, it would not be appropriate to compare the relative disparity for one objective expressed in terms of favorable events with the relative disparity for another objective expressed in terms of adverse events or to compare changes in disparity for these objectives over time.

Given the desire to compare disparity across the Healthy People 2010 objectives, a single approach —expression of all objectives in either favorable or adverse terms—was needed. A decision was made to measure disparity in terms of adverse events. Objectives that are stated in terms of favorable events (increase desired) are expressed in terms of adverse events (reduction desired) when measures of disparity are computed. The objectives themselves are not restated or changed.

Statistics for measuring health disparity

Pairwise statistics are being used to monitor progress toward the elimination of disparity for individual groups (compared to the best group rate) for all characteristics. For characteristics with three or more groups (e.g., race and ethnicity, education level, and income level), summary statistics are also used. A detailed description of the statistics and techniques employed to measure disparities in *Healthy People 2010* is provided in [2–4]. The key pairwise and summary statistics being used for *Healthy People 2010* disparity analyses are presented below.

Pairwise statistics

The percent difference is used to quantify disparities between the best group rate and another group rate. The percent difference is computed as:

$$\text{Percent difference} = \frac{R_\mathrm{i} - R_\mathrm{B}}{R_\mathrm{B}} \times 100,$$

where R_B is the best group rate for a particular characteristic and R_i is the rate for any other group of interest for a particular characteristic. For example, racial and ethnic health disparities are measured as the percent difference between the best racial and ethnic group rate and each of the other racial and ethnic group rates. See [3,4] for methods used in assessing the statistical significance of the percent difference.

Summary statistics (characteristics with three or more groups)

The index of disparity is used to determine whether the overall disparity from the best group rate is increasing or decreasing for the set of groups that make up a characteristic. It represents the *average percent difference* between the individual groups and the best group for the characteristic. The index of disparity is calculated as:

$$\text{Index of disparity} = \frac{\sum_{\mathrm{i=1}}^{\mathrm{n-1}} \mathrm{PD_i}}{\mathrm{n} - 1},$$

where PD_i is the percent difference from the “best” group rate for each of the groups of interest *(i)*, and *n-1* is the number of groups minus one. Because the percent difference is calculated with the best group rate as the reference point, the number of comparisons is equal to the number of groups minus one. See [4] for methods used in assessing the statistical significance of the index of disparity.

Assessing changes over time

When data beyond the baseline are available for an objective, the change in disparity from the baseline to the most recent data point can be measured. For pairwise comparisons, changes in disparity over time are measured by subtracting the percent difference from the best group rate at the baseline from the percent difference from the best group rate at the most recent data point. The change is expressed in percentage points: positive differences represent an increase in disparity and negative differences represent a decrease in disparity. Similarly, for comparisons involving three or more groups, the change over time is calculated by subtracting the index of disparity at the baseline from the index of disparity at the most recent data point. This statistic can be used to make comparisons over time only when data are available for the same groups defined in the same way at the baseline and at the most recent data point.

When standard errors for the constituent rates are available, the statistical significance of the change in the percent difference or the index of disparity over time can be evaluated [3,4]. When standard errors for the constituent rates are not available, the statistical significance of the change in the percent difference or the index of disparity over time cannot be evaluated.

References

1. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.
2. U.S. Department of Health and Human Services. Healthy People 2010 Midcourse Review. Washington, D.C.: U.S. Government Printing Office. 2006.
3. National Center for Health Statistics. Healthy People 2010 Final Review. Hyattsville, MD. 2011.
4. Keppel KG, Pearcy JN, Klein RJ. Measuring progress in Healthy People 2010. Healthy People Statistical Notes, no. 25. Hyattsville, MD: National Center for Health Statistics. September 2004. Available from: www.cdc.gov/nchs/data/statnt/statnt25.pdf.
5. Klein RJ, Proctor SE, Boudreault MA, Turczyn KM. Healthy People 2010 criteria for data suppression. Healthy People Statistical Notes, no. 24. Hyattsville, MD: National Center for Health Statistics. June 2002. Available from: www.cdc.gov/nchs/data/statnt/statnt24.pdf.
6. Keppel KG, Pearcy JN. Measuring relative disparities in terms of adverse outcomes. J Public Health Manag Pract 11(6). 2005.

4. Target Setting and Assessing Progress for Measurable Objectives

Target-Setting Methods

One of the three overarching goals for the Healthy People 2000 prevention initiative was to reduce health disparities among Americans [1]. The framework of Healthy People 2010 has taken this a step further by proposing to "eliminate health disparities" as one of the two overarching goals for the next decade.

To support the goal of eliminating health disparities, a single national target that is applicable to all selected populations has been set for each measurable, population-based objective. Three guiding principles were used in setting targets for the measurable, population-based objectives:

- For objectives that address health services and protection (for example, access to prenatal care and health insurance coverage) the targets have been set so that there is an improvement for all racial/ethnic segments of the population; that is, the targets are set to "better than the best" racial/ethnic subgroup shown for the objective. Data points for at least two population groups under the race and ethnicity category are needed to use "better than the best" as the target-setting method.
- For objectives that can be influenced in the short term by policy decisions, lifestyle choices, and behaviors (for example, physical activity, smoking, suicide, alcohol-related motor vehicle deaths), the target setting method is also "better than the best" group.
- For objectives that are unlikely to achieve an equal health outcome in the next decade, regardless of the level of investment (for example, occupational exposure and resultant lung cancer), the target represents an improvement for a substantial proportion of the population and is regarded as a minimum acceptable level. Implicit in setting targets for these objectives is the recognition that population groups with baseline rates already better than the identified target should continue to improve.

Beyond this general guidance, the exact target levels were determined by the Focus Area workgroups that developed the objectives. The workgroups used various methods for arriving at the target levels, including retention of the year 2000 target, computation of a statistical regression using current rates to project a target, knowledge of the programs currently in place and expected change, and expert judgment.

The following target-setting methods have been used:

- Better than the best
- _____ percent improvement
- "Total coverage" or "Total elimination" (for targets like 100%, 0%, all states, etc.)
- Consistent with _________ (another national program, for example, national education goals)
- Retention of the Healthy People 2000 target

The specific target-setting method is described under each objective in *Healthy People 2010* [2].

Baseline revisions

Targets were adjusted over the course of the decade for those objectives for which a change was made to the total population baseline data point after the publication of *Healthy People 2010* [2]. Baseline data were changed for a variety of reasons including revisions in methodology, survey questions, baseline year, and population denominators. Baseline data for several objectives were revised to accommodate updated public health recommendations. In several cases, baseline data were revised because the previously published data were based on preliminary analyses. Target revisions were not made in cases in which the baseline data for a select population had changed but data for the total population were unchanged.

The method used to adjust the target for an objective with a revised baseline for the total population depended on the original target-setting method outlined in *Healthy People 2010* [2].

- Targets based on "better than the best" racial/ethnic subgroup were revised using the same percent improvement from the racial/ethnic group with the "best" rate as was computed for the original target.
- Targets based on percent improvement were revised using the original percent improvement.
- Targets based on total elimination, total coverage, or consistent with another program were not revised.

Developmental objectives

The target-setting method for developmental objectives that became measurable during the 2005 midcourse review took into account the reduced time period (approximately five years) to achieve the target. The recommended method for setting targets for previously developmental population-based objectives was one unit "better than the best" racial and ethnic group, using the units in which the objective is measured. If the objective was measured in whole numbers, for example, the unit was 1. The "better than the best" approach represents the minimally acceptable improvement for all racial/ethnic groups. (Note that the original target-setting for population-based objectives was not limited to one unit "better than the best" racial/ethnic group.)

The Focus Area workgroups also had the option of proposing a target setting method greater or less than one unit "better than the best" population group (e.g., a specified percent improvement, etc.) subject to approval by the Assistant Secretary for Health.

Non-population-based developmental objectives that became measurable during the 2005 midcourse review set targets consistent with similar objectives in the Focus Area. Targets for these objectives could also take into consideration the reduced period to achieve the target.

Assessing Progress

Progress is assessed by the movement from the baseline value toward or away from the target. This is determined by a formula which indicates the change between the baseline and most recent data as a percent of the total change sought. The formula for assessing progress is as follows:

$$\text{Percent of targeted change achieved} = \frac{\text{Final value} - \text{Baseline value}}{\text{Healthy People 2010 target} - \text{Baseline value}} \times 100.$$

A number of objective headings contain multiple objectives. Progress is assessed separately for each objective, even if they have a common heading. Whenever possible, assessment of progress should consider the standard errors associated with the data (see Section 11: Variability of Estimates).

References

1. U.S. Department of Health and Human Services, Public Health Service. Healthy People 2000: National Health Promotion and Disease Prevention Objectives. Washington, D.C.: U.S. Government Printing Office. 1991.
2. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.

5. Population Estimates

Healthy People 2010 uses population estimates from the U.S. Census Bureau to calculate morbidity and mortality rates for many of the objectives [1–3]. Every 10 years, the U.S. Bureau of the Census (Census Bureau) conducts a full census of the resident population of the United States, Puerto Rico, and U.S. territories and collects data on sex, race, age, and marital status; the estimates produced represent the U.S. population as of April 1 of the census year. More detailed data on education, housing, occupation, income, and other information are also collected from a representative sample of the population (about 17 percent of the total population) [4].

The increasing diversity of the population has necessitated modification of the way race data are collected. In both the 1980 and 1990 censuses, a substantial number of persons did not specify a racial group that could be classified as any of the categories on the census form (white, black, American Indian, Eskimo, Aleut, Asian, or Pacific Islander) [5]. In 1980, the number of persons of "other" race was nearly seven million; in 1990 it was almost 10 million. In both censuses, the majority of these persons were of Hispanic origin (based on response to a separate question on the form), and many wrote in their Hispanic origin, or Hispanic origin type (for example, Mexican, Puerto Rican) as their race.

The Census Bureau presents population data by race in two different ways. In decennial census publications, persons of unspecified race are maintained in the single category of "other." For the purpose of providing comparable denominator data to other Federal and non-Federal data users, in both 1980 and 1990, the Census Bureau produced another set of population estimates for census years; in these population estimates, persons of unspecified race were allocated to one of the four tabulated racial groups (white, black, American Indian or Alaska Native, Asian or Pacific Islander), based on their response to the Hispanic question. These four race categories conformed to the Office of Management and Budget (OMB) Directive 15, "Race and Ethnic Standards for Federal Statistics and Administrative Reporting" and were more consistent with the race categories used in most major data systems, including vital statistics [6,7]. The postcensal and intercensal population estimates described below were based on these "OMB-consistent" populations.

In 1997, OMB issued "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity," which supersedes the 1977 Statistical Policy Directive 15 [8]. Both documents specify rules for the collection, tabulation, and presentation of race and ethnicity data within the Federal statistical system. The 1977 standards required Federal agencies to report race-specific tabulations using four single-race categories, namely, White, Black, American Indian or Alaska Native, and Asian or Pacific Islander. The 1997 revision incorporated two major changes designed to reflect the changing racial and ethnic profile of the United States. First, the 1997 revision increased from four to five the minimum set of categories to be used by Federal agencies for identification of race. As in the past, these categories represent a social-political construct and are not anthropologically or biologically based.

The five categories for race specified in the 1997 standards are: American Indian or Alaska Native; Asian; Black or African American; Native Hawaiian or Other Pacific Islander; and White. Second, the revised standards add the requirement that Federal data collection programs allow respondents to select one or more race categories when responding to a query on their racial identity. This provision means that there are potentially 31 race groups, depending on whether an individual selects one, two, three, four, or all five of the race categories. Collection of additional detail on race or ethnicity is permitted so long as the additional categories can be aggregated into the minimum categories.

In recent years, data systems have been revising their collection and tabulation procedures to comply with the 1997 standards. Some data systems implemented the new standards between 1999 and 2003, while others are still in the process of planning for or implementing the new standards. Therefore, templates for race and Hispanic origin vary across objectives. In addition, the data systems used to track the population-based Healthy People 2010 objectives in may not provide data for all of these domains and subgroups. However, some data systems provide data for additional subgroups, for example, Hispanic origin subgroups such as Cuban, Mexican American, and Puerto Rican [9].

During the transition to full implementation of the 1997 standards, two different standards for the collection of race and ethnicity data are being used, creating incomparability across data systems. Further, within a given data system, the change in the race standards results in incomparability across time, thus making it difficult to perform trend analyses. The OMB recognized that approaches to make data collected under the 1997 standards comparable to data collected under the 1977 standards would be needed. Therefore, the OMB issued "Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity" [10]. The guidance document contains a detailed discussion of bridging methods. (Bridging is a mechanism for collapsing the multiple-race- group population counts into single-race-group counts.) The vast majority of the Healthy People 2010 major data systems (these are data systems tracking five or more objectives) have implemented the 1997 Standards.

The National Vital Statistics System is one of the Healthy People 2010 major data sources that has not yet fully implemented the 1997 standard. This is because vital rates are based on information obtained from vital records collected through the state-based Vital Statistics Cooperative Program (numerators) and population estimates derived from the U.S. Census (denominators). Implementation of the 1997 standards within the Vital Statistics Cooperative System started in 2003, on an individual state basis, and is proceeding slowly as states implement revised birth and death certificates which incorporate the revised OMB standards. As of 2007, most states had not revised the race question on their birth or death certificates and were still collecting race data using the 1977 race categories. Other major data systems (such as National Survey of Family Growth, the HIV Surveillance System, STD Surveillance System) have implemented the 1997 Standard, but are not reporting data for all the domains. One of the reasons could be attributed to sample size of the race categories.

Postcensal Population Estimates

National population estimates for the years after the decennial census (postcensal estimates) are calculated using the decennial census as the base population and adjusting those counts using the following measures of population change: births and deaths (provided by the National Center for Health Statistics), immigration data (provided by the U.S. Immigration and Naturalization Service), data on the movement of Armed Forces personnel (from the U.S. Department of Defense [DoD]), movement between Puerto Rico and the U.S. mainland (from Puerto Rico Planning Board), and movement of Federal employees abroad (from the Office of Personnel Management and DoD). These estimates reflect the U.S. population as of July 1 of each year. Postcensal estimates for state and county populations are also calculated using these data, as well as data from the Internal Revenue Service and State departments of education. Postcensal estimates become less accurate as the date of the estimates moves farther from the date of the census [11].

The population estimates in the postcensal years are based on the April 1, 2000, resident population as enumerated by the Census Bureau. They result from bridging the 31 race categories used in Census 2000, as specified in the 1997 OMB standards, to the four race groups specified under the 1977 OMB standards. The bridged-race postcensal estimates were prepared by the Census Bureau’s Population Estimates Program under a collaborative agreement with the National Center for Health Statistics [12].

Intercensal Estimates

After each decennial population census, intercensal estimates for the preceding decade are calculated to replace postcensal estimates. These estimates reflect the population as of July 1. Intercensal estimates are more accurate than postcensal estimates because they incorporate data from the enumerations at the beginning and end of the decade. The intercensal estimates for the 1980s were used to revise some of the baselines for mortality objectives in Healthy People 2000; these were published in the *Healthy People 2000 Midcourse Review and 1995 Revisions* [13]. Intercensal estimates for 1991–99, incorporating the bridged 2000 Census data, were produced by the Census Bureau under a collaborative agreement with NCHS [14]. These estimates have been used to recalculate some of the morbidity and mortality data for Healthy People 2010 baselines and tracking data for data years prior to the year 2000.

Population Undercounts

Some subgroups of the population (including some racial, ethnic, and age groups) are less likely than other groups to be completely enumerated in the decennial census. The undercounts of these groups lower the denominators and result in higher morbidity and mortality rates for these populations [15,16]. The Census Bureau makes estimates of net census undercount for the total, white, and black populations by age. These estimates are then used to weight the populations used by most of the national health surveys, including National Health Interview Survey, the National Health and Nutrition Examination Survey, the National Survey of Family Growth, and the National Health Care Surveys.

The National Vital Statistics System (Mortality and Natality) use population denominators that are not adjusted for net census undercount.

Target Populations

Several types of target populations are used for Healthy People 2010 objectives*:*

Resident Population

The resident population includes all persons whose usual place of residence is in one of the 50 states or the District of Columbia, including Armed Forces personnel stationed in the United States. The resident population is usually the denominator when calculating birth and death rates and rates of new cases of disease. The resident population is also the denominator for selected population-based rates that use numerator data from the National Nursing Home Survey.

Civilian Population

The civilian population is the resident population, excluding members of the Armed Forces (although their family members are included). The civilian population is the denominator for other Healthy People 2010 data sources, such as the National Hospital Discharge Survey.

Civilian, Noninstitutionalized Population

The civilian, noninstitutionalized population is the civilian population not residing in institutions (for example, correctional facilities, psychiatric hospitals, and nursing homes). This population is the denominator for rates from Healthy People data sources such as the National Ambulatory Medical Care Survey and the National Hospital Ambulatory Care Survey. This population is also used in the weighting procedure to produce national estimates from health surveys such as National Health Iinterview Survey, the National Health and Nutrition Examination Survey, and the National Household Survey on Drug Abuse.

Details on the specific populations targeted for each major Healthy People 2010 data system can be found in the data source tables included in *THP, Part C*. The objective operational definitions shown in *Tracking Healthy People, Part B* indicate the population covered by each objective, if applicable.

References

1. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.
2. U.S. Department of Health and Human Services. Healthy People 2010 Midcourse Review. Washington, D.C.: U.S. Government Printing Office. 2006.
3. National Center for Health Statistics. Healthy People 2010 Final Review. Hyattsville, MD. 2011.
4. Detailed information on the methodology used to collect census data is available from the U.S. Bureau of the Census Web site: http://www.census.gov.
5. U.S. Bureau of the Census. U.S. population estimates, by age, gender, race and Hispanic origin, 1980–91. Current Population Reports P-25(1095). Washington, D.C.: U.S. Department of Commerce. 1992.
6. Office of Management and Budget. Directive No. 15. Race and ethnic standards for Federal statistics and administrative reporting. Statistical Policy Handbook. Washington, D.C.: Office of Management and Budget. 1978.
7. National Center for Health Statistics. Vital Statistics of the United States, 1992, Vol. I, Natality, Technical Appendix. Hyattsville, MD: National Center for Health Statistics. 1996. Available from: http://www.cdc.gov/nchs/data/vsus/nat92_1.pdf.
8. Office of Management and Budget. Standards for maintaining, collecting, and presenting Federal data on race and ethnicity. Federal Register 62: 58781–90. 1997.
9. Keppel KG, Pearcy JN, Klein RJ. Measuring progress in Healthy People 2010. Healthy People Statistical Notes, no. 25. Hyattsville, MD: National Center for Health Statistics. September 2004. Available from: www.cdc.gov/nchs/data/statnt/statnt25.pdf.
10. Provisional Guidance on the implementation of the 1997 Standards for Federal Data on Race and Ethnicity. Washington, D.C.: Executive Office of the President, Office of Management and Budget. December 15, 2000.
11. Byerly ER, Deardorff K. National and State Population Estimates: 1990 to 1994. Current Population Reports P25-1127. Washington, D.C.: U.S. Bureau of the Census. U.S. Government Printing Office, 1995.
12. Ingram DD, Parker JD, Schenker N, et al. United States Census 2000 population with bridged race categories. National Center for Health Statistics. Vital Health Stat 2(135). Hyattsville, MD. 2003. Available from: http://www.cdc.gov/nchs/data/series/sr_02/sr02_135.pdf.
13. U.S. Department of Health and Human Services, Public Health Service. Healthy People 2000 Midcourse Review and 1995 Revisions. Washington, D.C.: U.S. Government Printing Office. 1995.
14. National Center for Health Statistics. Estimates of the July 1, 1991–1999 United States resident population by State and county, age, sex, race, and Hispanic origin, prepared under a collaborative arrangement with the U.S. Census Bureau. Released April 15, 2003. Available from: www.cdc.gov/nchs/about/major/dvs/popbridge/popbridge.htm.
15. National Center for Health Statistics. Vital Statistics of the United States, 1992, Vol. II Mortality, Part A, Technical Appendix. Hyattsville, MD: National Center for Health Statistics. 1996. Available from: http://www.cdc.gov/nchs/data/vsus/mort92_2a.pdf.
16. Rosenberg HM, Maurer JD, Sorlie PD, et al. Quality of death rates by race and Hispanic origin: A Summary of current research, 1999. National Center for Health Statistics. Vital Health Stat 2(128), Hyattsville, MD, 1999. Available from: http://www.cdc.gov/nchs/data/series/sr_02/sr02_128.pdf.

6. Population Template

Minimum Template

During the review of the September 1998 *Healthy People 2010 Draft for Public Comment*, the need for greater consistency in tracking population groups became apparent [1]. To address this issue, a minimum template for all population-based Healthy People 2010 objectives was adopted. Population-based objectives may show more detailed and additional breakouts if appropriate.

This minimum select population template applies to most measurable population-based objectives and is also applied to developmental population-based objectives (see Section 4, Developmental Objectives) as data become available. The template does not apply to non-population-based objectives such as those that measure schools, worksites, or states. Because of problems in interpreting risk, the template is also not shown for population-based measurable objectives that are tracked using counts of events rather than rates or percents.

The current minimum template for all population-based objectives is:

- Race:
  - American Indian or Alaska Native only
  - Asian only
  - Black or African American only
  - Native Hawaiian or Other Pacific Islander only
  - White only
  - Two or more races
    - American Indian and Alaska Native; White
    - Black or African American; White
- Hispanic Origin and Race:
  - Hispanic or Latino
  - Not Hispanic or Latino
    - Black or African American only
    - White only
- Gender:
  - Female
  - Male
- Socioeconomic status:

Family income level	or	Education level
Poor		Less than high school
Near poor		High school graduate
Middle/high income		At least some college

Additional subgroups are included for specific objectives, including: geographic location (urban/rural), health insurance status, disability status, chronic disease status, sexual orientation, and specific age groups. These subgroups are defined elsewhere in this publication.

The racial categories differ from those used with Healthy People 2000 and in the original *Healthy People 2010* publication [2]. (See Race and Hispanic Origin, below.) In the original *Healthy People 2010* publication the race template did not specify 'only' for single race groups and it did not include 'Two or more races' categories [2]. The minimum race template in the original publication was:

- Race:
  - American Indian or Alaska Native
  - Asian or Pacific Islander
    - Asian
    - Native Hawaiian or Other Pacific Islander
  - Black or African American
  - White
- Hispanic Origin and Race:
  - Hispanic or Latino
  - Not Hispanic or Latino
    - Black or African American
    - White

The groups listed under most headings (race, Hispanic origin, sex, and income) in the minimum template are comprehensive; that is, they are intended to sum to the population (excluding "unknowns") tracked by the objective. For example, the three groups under income equal the total population tracked by the objective. The exception is the education category, which is limited to people of a minimum age or, in some cases, a maximum age (see Socioeconomic Status discussion below). The groups listed under the subheading "Not Hispanic" are not inclusive.

If data are not provided for a group, this is indicated by one of four statements: data have been collected but have not yet been analyzed (DNA), data are not collected by the data system used to track the objective (DNC); data do not meet the criteria for statistical reliability, data quality, or confidentiality (DSU); or the specific breakout is not applicable (NA). In cases where data for the entire template are not collected by the data system tracking the objective, a note to this effect will replace the template. For more information on statistical reliability, see Section 11: Variability of Estimates and Data Suppression.

Race and Hispanic Origin

OMB Classification

On October 30, 1997, the Office of Management and Budget (OMB) published "Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity" [5]. These standards revised the 1977 OMB Directive No. 15, "Race and Ethnic Standards for Federal Statistical Reporting [6]. The revised standards modified the Federal data collection policy, requiring Federal agencies to collect information that reflects the increasing diversity of our Nation's population.

The revised standards were used by the U.S. Census Bureau in the 2000 decennial census. Other Federal programs were required to incorporate them into household surveys, administrative forms and records, and other data collections by January 1, 2003. Section 5: Population Estimates, includes a discussion of issues related to the transition from OMB Directive 15 to the 1997 Standards.

OMB Directive No. 15 defined the basic racial and Hispanic origin categories for Federal statistics and program administrative reporting as American Indian or Alaska Native, Asian or Pacific Islander, Black, White, and Hispanic [6]. The 1997 policy requires agencies to offer respondents the option of selecting one or more of the following five racial categories:

- **American Indian or Alaska Native**. A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.
- **Asian**. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
- **Black or African American**. A person having origins in any of the black racial groups of Africa.
- **Native Hawaiian or Other Pacific Islander**. A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
- **White**. A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

These five categories are the minimum set for data on race for Federal statistics, program administrative reporting, and civil rights compliance reporting. The new standards explicitly do not include an "other race" category for data collection; however, an "other" category may be used for tabulating and data reporting [6].

The "some other race" option is increasingly being selected in the U.S. Census. In the 2010 census, of persons self-identifying their race as "some other race," 95 percent also identify themselves as Hispanic; 36.7 percent of Hispanics selected "some other race" as their sole racial identification [7].

The OMB standards require that at a minimum, the total number of persons identifying with more than one race be reported when data are available. It is stressed that this is a minimum; the presentation of detailed information on specific racial combinations subject to constraints of data reliability and confidentiality standards is preferred. Over time, this percentage may increase as those who identify with more than one race become aware of the opportunity to report more than one race group. In the 2010 census, 2.9% identified with two or more races [7].

The standards regarding Hispanic origin provide for the collection of data on whether or not a person is of “Hispanic or Latino” culture or origin. This category is defined as follows:

Hispanic or Latino. A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. The term, “Spanish origin,” can be used in addition to “Hispanic or Latino.” Persons of Hispanic origin may be of any race and persons in the various race groups may be of any origin.

To provide flexibility and to assure data quality, the new OMB guidelines recommend that a two-question format (separate race and Hispanic ethnicity questions) be used, especially when respondents can self-identify. When race and ethnicity are collected separately, ethnicity should be collected first. Most Healthy People data systems that use self-identification, such as the National Health Interview Survey and the National Health and Nutrition Examination Survey, use the two-question format. When self-identification is not feasible (for example, the National Notifiable Disease Surveillance System) or when there are overriding data collection considerations (for example, the Youth Risk Behavior Surveillance System prior to 2007), a combined race and ethnicity question can be used that includes a separate Hispanic category co-equal with the other (racial) categories. When a combined question is used, more than one entry (race and ethnicity or multiple races) is possible.

Misreporting Racial and Ethnic Data

Most health surveys and censuses obtain the self-reported race of the respondent. This is considered to be the most accurate representation of a person’s racial or ethnic background. However, some data systems cannot collect self-reported race or ethnicity. For example, the National Vital Statistics System mortality component collects information about the decedent from an informer. In other systems, such as those derived from hospital/patient care records, it is often unclear whether the information is self-reported. In these cases, race and ethnicity may be entered by someone else (for example, clerical staff or hospital personnel) based on observation or the report of proxy respondents. Several of these data systems are discussed below.

National Vital Statistics System (Mortality)

Death rates by race and Hispanic origin may be biased from misreporting of race and Hispanic origin in the numerator of the rates and misreporting and undercoverage in the denominator of the rates [8]. Numerator data are from the death certificate as reported by the funeral director based on information from an informant, usually a family member [9]. Denominator data, from surveys or the Census, is either self-reported or reported by a member of the household. Past studies comparing death certificate information with that from independent sources such as the Current Population Survey, indicate that the reporting of race on the death certificate is good for the white and black populations; however, the reporting of race and Hispanic origin for other groups may be seriously under-stated [10]. Additional problems, such as population undercounts (see Section 5: Population Estimates), affect population censuses and estimates [11]. As a consequence of the combined effect of numerator and denominator biases, it has been estimated that death rates for the white and black populations are overestimated by about one percent and five percent respectively. Death rates are underestimated for the American Indian or Alaska Native population by approximately 21 percent; for the Asian or Pacific Islander population by 11 percent; and for the Hispanic population by two percent [8]. These estimates are approximations; they do not take into account differential misreporting by age and sex among the race/ethnic origin groups.

An update to the report concluded that the reporting remains poor for the AIAN population, and that reporting for Hispanic and API populations is ‘reasonably good’. This report estimated that death rates are underestimated for the AIAN population by approximately 30%; for the API population by 7% and for the Hispanic population by 5% [12].

For *Healthy People 2010,* infant mortality rates for races and ethnic populations are based on linked files of infant deaths and live births [3,4,13]. These rates use the race of mother as self-reported on the birth certificate and, therefore, are not affected by the misreporting of race on the death certificate.

Patient Care Data

Systems that collect data from patient records such as the National Hospital Discharge Survey (NHDS), the National Ambulatory Medical Care Survey, the National Hospital Ambulatory Medical Care Survey, the National Notifiable Disease Surveillance System, the HIV/AIDS Surveillance System, also may misreport the race of individuals. It is often unclear how race and ethnicity are reported in these systems. The race and ethnicity of the patient may be reported by hospital or other medical care personnel by observation, by proxy report, or by the patient. Therefore, one must use information on race and ethnicity from these systems with caution.

Missing Data

In addition to the problems of misreporting race and ethnicity, the information on race reported by some data systems are often missing or incomplete. Some of these systems are described below. Specific information on the quality and completeness of reporting of race and ethnicity for the major Healthy People 2010 data systems is included, where available, in the data source description in *Tracking Healthy People, Part C.*

National Hospital Discharge Survey

Race is not reported in about 18 percent of NHDS records since data on race are not reported by many hospitals due to the omission of a race field on hospital discharge reporting forms [14]. More hospitals have automated their discharge systems in recent years and are currently using form UB-92 which does not require race reporting. A comparison of NHDS data with data on persons who reported being hospitalized in the National Health Interview Survey (NHIS) (NHIS data were adjusted to exclude hospitalizations of one day or less) indicated that underreporting for the white patients was about 22 percent in 1991; the difference for African Americans was negligible [14]. Hispanic origin is not reported for approximately 50 percent of the NHDS records [15].

National Ambulatory Medical Care Surveys

Race is not reported for about 10 to 32 percent of records in the National Ambulatory Medical Care Survey and the National Hospital Ambulatory Care Survey. However, missing values are imputed for both surveys [16–21].

National Notifiable Disease Surveillance System

Although staff in State health departments and the Centers for Disease Control and Prevention attempt to obtain complete demographic information associated with nationally notifiable cases of disease, some data (particularly for the variables of race and ethnicity) are not available for some cases of disease. Laws, regulations, and mandates for public health reporting (including specific data items that are reported) fall under the authority of individual states, and in some states, race and ethnicity may not be approved for reporting to the national level. Race and ethnicity data may also be unknown when cases are reported from a laboratory or when cases are reported as aggregate disease totals.

Socioeconomic Status

One of the three overarching goals for the Healthy People 2000 prevention initiative was to reduce health disparities among Americans [22]. Healthy People 2010 has taken this a step further, making the elimination of health disparities one of the two primary goals to be achieved by 2010. While disparities among racial and ethnic groups—especially between whites and blacks—received considerable attention over the last decade, differential health outcomes and access to social and health care resources often reflect differences in education, occupation, income, and wealth. Monitoring progress toward eliminating social and economic disparities in health requires improved collection and use of standardized data on the socioeconomic status of individuals.

Socioeconomic status (SES) may be represented by income, level of education, or type of occupation. *Healthy People 2010* uses education and income-related measures as primary measures of SES [2–4]. The following discussion presents data issues for income and education measures.

Income

Income is the most common measure of SES and is probably the most relevant to health policy formulation. Current income provides a direct measure of the quality of food, housing, leisure-time amenities, and health care an individual is able to acquire, as well as reflecting the relative position in society. However, income may fluctuate over time so that income received in a given year may not accurately reflect one's lifetime income stream or total wealth, the measures of resources more relevant to health. For example, elderly persons who have low incomes may also have accumulated assets that offset their need for a high annual income. Of particular importance in considering the relationship between income and health is the fact that income may be low because illness has limited the amount of income earned or prevented earning income entirely. The use of income as a measure of SES also involves more practical difficulties. In many heath surveys a substantial number of persons either do not know or refuse to report their incomes [23].

Family Income

The National Health Interview Survey (NHIS) and the National Health and Nutrition Examination Survey consider all persons within a household who are related to each other by blood, marriage, or adoption to constitute a family. Each member of a family is classified according to the total income of the family. Unrelated individuals are classified according to their own income. Since 1997, the NHIS has collected family income data for the calendar year prior to the interview (for example, 2003 family income data were based on 2002 calendar year information). Family income includes wages, salaries, rents from property, interest, dividends, profits and fees from their own businesses, pensions, and help from relatives. Family income data are used in the computation of the poverty level. To handle the problem of missing data on family income in the NHIS, multiple imputations were performed for survey years 1997-2008 (and each later year as data are available) with five sets of imputed values created to allow for the assessment of variability due to imputation. Family income was imputed for 24–34 percent of families from 1997–2008 [24]. A detailed description of the imputation procedure is available from: www.cdc.gov/nchs/ nhis/2008imputedincome.htm.

When income is selected for the template, poor, near poor, and middle/high income categories are used unless overridden by programmatic or data considerations (for example, eligibility for the Special Supplemental Nutrition Program for Women, Infants, and Children). In these special cases, the poverty categories appropriate for the program or system are used. For most health surveys, income is defined as money income before taxes and does not include the value of non-cash benefits such as food stamps, Medicare, Medicaid, public housing, and employer-provided fringe benefits.

Poverty Level

Converting income to poverty status adjusts for family size and inflation, facilitating comparisons among groups and over time. Poverty status measures family income relative to family size using the poverty thresholds developed by the U.S. Bureau of the Census, based on definitions originally developed by the Social Security Administration. These thresholds vary by family size and composition and are updated annually to reflect changes in the Consumer Price Index for all urban consumers. Families or individuals with income below their appropriate thresholds are classified as below the poverty level.

Focusing simply on the dichotomy of “above” versus “below” poverty, however, obscures the full gradient of inequalities in income distribution and in health. Understanding burden across the income gradient provides information useful for potential eligibility expansions or other programmatic modifications. In *Healthy People 2010*, the three categories of family level income that are primarily used are [3,4]:

- Poor (below the Federal poverty level)
- Near poor (100-199% of the Federal poverty level)
- Middle and high income (200% or more of the Federal poverty level).

For a family of four, the average Federal poverty level weighted for family composition was $18,810 in 2003. In 2008 the weighted average threshold for a family of four was $22,025. Table 1 shows the 2003 and 2008 poverty thresholds by size of family and number of related children aged under 18 years.

Table 1. Poverty Thresholds in 2003 and 2008, by Size of Family and Number of Related Children Aged Under 18 Years

2003 Poverty Thresholds

Size of Family Unit	**Related Children Aged Under 18 Years**								
	None	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
One person									
Aged under 65 years	9,573								
Aged 65 years and over	8.825								
Two persons									
Householder aged under 65 years	12,321	12,682							
Householder aged 65 years and over	11,122	12,634							
Three persons	14,393	14,810	14,824						
Four persons	18,979	19,289	18,660	18,725					
Five persons	22,887	23,220	22,509	21,959	21,623				
Six persons	26,324	26,429	25,884	25,362	24,586	24,126			
Seven persons	30,280	30,479	29,827	29,372	28,526	27,538	26,454		
Eight persons	33,876	34,175	33,560	33,021	32,256	31,285	30,275	30,019	
Nine persons or more	40,751	40,948	40,404	39,947	39,196	38,163	37,229	26,998	35,572

2008 Poverty Thresholds

Size of Family Unit	Related Children Aged Under 18 years								
	None	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8+**
One person									
Aged under 65 years	11,201								
Aged 65 years and over	10,326								
Two people									
Householder aged under 65 years	14,417	14,840							
Householder aged 65 years and over	13,014	14,784							
Three people	16,841	17,330	17,346						
Four people	22,207	22,570	21,834	21,910					
Five people	26,781	27,170	26,338	25,694	25,301				
Six people	30,803	30,925	30,288	29,677	28,769	28,230			
Seven people	35,442	35,664	34,901	34,369	33,379	32,223	30,955		
Eight people	39,640	39,990	39,270	38,639	37,744	36,608	35,426	35,125	
Nine people or more	47,684	47,915	47,278	46,743	45,864	44,656	43,563	43,292	41,624

Note: Numbers represent income in U.S. dollars.
Sources: U.S. Bureau of the Census, Current Population Survey, 2004 Annual Social and Economic Supplement and http://www.census.gov/hhes/www/poverty/data/threshld/.

In addition to the limitations discussed for income, converting income to poverty status introduces other issues that need to be considered. If income data are collected by selecting an appropriate income category, rather than giving the actual dollar amount, then the conversion to poverty status must be performed using category means or medians and will thus result in some misclassification. An alternative is to convert the chosen income category to poverty status using the range of incomes; in this case some of the income categories will include a range of incomes that are 'borderline' because some of the incomes in the range will fall below the poverty threshold while others in the same income category will fall above the poverty threshold.

Education

Education is frequently used as the measure of SES in presentations of health data. There are several reasons for this preference. Education is generally more completely reported than income; usually 95 percent or more of respondents report their attained level of education. Unlike occupation, all adults may be characterized by their education level. Education, unlike income or occupation, remains fixed for most people after the age of 25 and usually is not influenced by health. In addition, education is highly related to both income and occupation.

Education cannot be used to characterize the socioeconomic position of children (except through the educational level of parents or head of household). The average education level of the U.S. population has increased steadily over time, complicating comparisons across age groups. Between 1975 and 2010, the educational attainment of persons aged 25 to 29 years completing high school rose from 83 to 89 percent; the percentage with some college rose from 42 to 61 percent; and the percentage with 4 or more years of college rose from 22 to 32 percent [25].

Educational attainment is typically measured either by the number of years of education the individual has completed or by the highest credential received. The categories for educational attainment that are primarily used in *Healthy People 2010* are [3,4]:

- Less than high school (persons with less than 12 years of schooling or no high school diploma)
- High school graduate [persons with either 12 years of schooling, a high school diploma, or Certificate of General Educational Development (GED)]
- At least some college (persons with a high school diploma or GED and 13 or more years of schooling).

In general, data on educational attainment are presented for ages beginning with 25 years, consistent with guidance given by the U.S. Bureau of the Census. However, objectives using different data systems may have different age groups for the education variable. The actual ages that are used to calculate educational attainment for some of the major Healthy People 2010 data systems are shown in Table 2. Because of the requirements of the different data systems, the age groups used to calculate educational attainment for an objective may differ from the age groups used to report the data for other select populations and the overall measure of the same objective. For clarity, each Healthy People 2010 objective states the age groups used to measure the levels in the educational attainment category. Caution must be used in comparing the data by educational attainment with data for the main objective and other select populations.

Healthy People 2010 baseline education data for the mortality objectives are based on reports from 46 states and the District of Columbia. Mortality statistics do not report data by education for the elderly population (aged 65 years and over) because the percentage with “education not stated” is higher for this group and because of possible bias due to misreporting of education on the death certificate. The death rate for high school graduates (12 years of education) is generally overstated because there is a tendency for some people who did not graduate from high school to be reported as high school graduates on the death certificate; by extension, the death rate for the group with less than 12 years of education tends to be understated [26].

Beginning in 2003, education data from the National Vital Statistics System have been suppressed. The educational attainment item was changed in the new U.S. Standard Certificates for Births, Deaths, and Fetal Death in 2003 to be consistent with the U.S. Census Bureau data and to improve the ability to identify specific types of educational degrees. Many states, however, are still using the 1989 version of the U.S. Standard Certificates, which focuses on highest school grade completed. As a result, educational attainment data collected using the 2003 version are not comparable with data collected using the 1989 version [27–30].

Table 2. Healthy People 2010 Data Systems and Ages Used to Report Educational Attainment

Data System	**Ages Used To Report Educational Attainment**
Behavioral Risk Factor Surveillance System (BRFSS)	25 years and over
National Health and Nutrition Examination Survey (NHANES)	25 years and over (unless otherwise noted)
National Health Interview Survey (NHIS)	25 years and over
National Survey of Family Growth (NSFG)	20–44 years
National Vital Statistics System—Mortality (NVSS-M)	25–64 years
National Vital Statistics System—Natality and Linked (NVSS-N and NVSS-L)	20 years and over
National Survey on Drug Use and Health (NSDUH)	18 years and over

Other Population Groups

Several other groups were considered for inclusion in the minimum set of select populations but were left to the discretion of the Public Health Service agencies responsible for each Healthy People 2010 Focus Area to include under specific objectives where appropriate. These groups included urban/rural residence, health insurance status, disability status, age, sexual orientation, the institutionalized population, and immigrant status, some of which are discussed in greater detail below. Some objectives also include select populations of persons with specific conditions—such as persons with diabetes, persons with hypertension, and persons with arthritis.

Urbanization

Urban residence in *Healthy People 2010* is specified as either residing within or outside a metropolitan statistical area or residing within or outside an urbanized area or urban place (called "urban" in the template) as designated by the U.S. Census Bureau [3,4].

Urban

Urban residence is defined as people living within the boundaries of an urbanized area and the urban portion of places outside an urbanized area that have a decennial census population of 2,500 or more. An urbanized area is an area consisting of a central place(s) and adjacent urban fringe that together have a minimum residential population of at least 50,000 people and generally an overall population density of at least 1,000 people per square mile of land area. The U.S. Census Bureau uses published criteria to determine the qualification and boundaries of urbanized areas. For more information see the Census Bureau Web site, accessible from http://www.census.gov/geo/www/tiger/glossary.htm.

Metropolitan Statistical Areas (MSA's)

Metropolitan statistical areas (MSAs) are established by the U.S. Office of Management and Budget. The MSA standards are revised before each decennial census. When census data become available, the standards are applied to define the actual MSAs. An MSA is a county or group of contiguous counties that contains at least one urbanized area of 50,000 or more population. In addition to the county or counties that contain all or part of the urbanized area, an MSA may contain other counties that are metropolitan in character and that are economically and socially integrated with the main city. In New England, cities and towns, rather than counties, are used to define MSAs. Counties that are not within an MSA are considered to be nonmetropolitan. For current and historical definitions of MSAs, see [31].

Health Insurance Status

The health insurance status template applies only to persons aged under 65 years. Those aged 65 years and over are considered to be covered by Medicare. Respondents are identified as having health insurance if they are covered by either private or public health plans. Private insurance includes fee-for-service plans, single service hospital plans, and coverage by health maintenance organizations. Public insurance includes Medicaid or other public assistance, Temporary Assistance for Needy Families, Supplementary Security Income, Medicare, or military health plan coverage.

Disability

In 1980, the World Health Organization published the first version of the International Classification of Impairments, Disabilities, and Handicaps (ICIDH) as a classification of the "consequences of disease" [32]. The International Classification of Functioning, Disability, and Health (ICF) was published in 2001 [33].

According to IFC, components of disability include:

- Impairments to body functions
- Impairments to body structures
- Limitations to participation in activities with or without assistance or the use of assistive devices
- Barriers and facilitators which make up the physical, social and attitudinal environment (environmental factors).

The major sources of national data on people with disabilities include:

- Decennial Census
- Survey of Income and Program Participation
- National Health Interview Survey (NHIS)
- National Health and Nutrition Examination Survey (NHANES)
- Medical Expenditure Panel Survey
- Current Population Survey.

In *Healthy People 2010*, the major sources of disability data are the NHIS and NHANES for national data and the Behavioral Risk Factor Surveillance System (BRFSS) for state-level data [3,4]. The NHIS has several variables that can be used to operationally define disability status, including limitation of activity, restriction of participation (bed days, work-loss days, school-loss days), and assessed health status [34]. The BRFSS also collects information on health-related quality of life, limitation of activity, and self-assessed health status.

Disability is operationally defined in a number of different ways for program purposes and for analytic and research purposes, depending on the data collected by the data systems. In *Healthy People 2010*, disability is primarily defined using information on activity limitation or the use of special equipment [3,4]. The definitions used by the NHIS, BRFSS, and NHANES are described below.

In the 1997 NHIS, a person is classified as having a disability if a "yes" response was obtained to <u>any</u> of the age-appropriate limitation questions or to the use of special equipment. [See the operational definition for the denominators used for objectives 6-2 (children) and 6-3 (adults) in *Tracking Healthy People, Part B,* for the specific questions used from the 1997 NHIS.] For NHIS data prior to 1997, the special equipment questions were not asked, so persons are categorized in the templates as "with activity limitation" rather than "with disabilities."

State data are available from the BRFSS telephone surveys. For *Healthy People 2010*, using 2001 BRFSS data, people answering "yes" to *any* of the following questions define adults aged 18 years and over with disabilities [3,4]:

- Are you limited in any way in any activities because of any impairment or health problem?
- Do you now have any health problem that requires you to use special equipment, such as a cane, a wheelchair, a special bed, or a special telephone?

In 2003 these questions were revised to:

- Are you limited in any way in any activities because of physical, mental, or emotional problems?
- Do you now have any health problem that requires you to use special equipment, such as a cane, a wheelchair, a special bed, or a special telephone? Include occational use or use in certain circumstances.

The 1999–2002 NHANES is also used to classify persons with disability. Persons aged 20 years and over were identified by NHANES as having a disability if they met any of the following criteria:

- Unable to work at a job or business because of a physical, mental, or emotional problem
- Limited because of difficulty remembering or because of periods of confusion
- Limited in any activity because of a physical, mental or emotional problem
- Uses special equipment, such as a cane, a wheelchair a special bed, or a special telephone.

Disability data from the NHANES III were limited to the second phase (1991–94) and are calculated only for people aged 20 years and over. People are classified as having a disability if a "yes" response was obtained to *any* of the following questions:

- Are you limited in the kind or amount of *work* you can do because of any impairment or health problem?
- Are you limited in the kind or amount of *housework* you can do because of any impairment or health problem?
- Are you limited in any way in any *activities* because of any impairment or health problem?
- Do you usually use any device to help you get around such as a cane, wheelchair, crutches or walker?

The Centers for Disease Control and Prevention proposed that a standardized set of questions on disability status be developed. There are two questions that were proposed: 1) Are you limited in any way in any activities because of physical, mental or emotional problems? and/or 2) Do you now have any health problem that requires you to use special equipment, such as a cane, a wheelchair, a special bed or a special telephone? As these constructs (regardless of minor differences in wording) are adopted by the data systems, the data produced from them are being incorporated into the Healthy People 2010 objectives that specifically identify people with disabilities using this standard definition. This presents the opportunity to have a standard definition of people with disabilities that can be used across data systems and geographic levels. Healthy People 2010 objective 6-1 tracks the incorporation of this standard definition in major data systems used to monitor the Healthy People 2010 objectives. As of 2009, 33 percent of relevant data systems had adopted the standard questions.

To a large extent, disability measures are related to the generation of many summary measures discussed in the goals section of *Healthy People 2010* [2]. Summary measures of health generally combine information on mortality and health into a single measure. Many of these summary measures use variables that directly relate to disability status to generate the health component (often referred to as health-related quality of life) of the measure. Because of this, disability measures have importance beyond the assessment of the disability status of a population.

Age

Age is not included in the minimum template because showing inclusive age categories would add considerable complexity to the minimum set. Furthermore, age is often stated in the objective (for example, mammograms for females aged 40 years and over) and many objectives are relevant only for a subset of age groups. Age-specific select populations are added to objectives where needed and may not be inclusive of the total population. For example, age-specific measures for the elderly, adolescents, or children have been added to some objectives without adding other groups, although showing inclusive age breakouts, if relevant, is preferred.

References

1. U.S. Department of Health and Human Services (DHHS). Healthy People 2010 Objectives: Draft for Public Comment. Washington, D.C.: DHHS, Office of Public Health and Science, 1998.
2. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.
3. U.S. Department of Health and Human Services. Healthy People 2010 Midcourse Review. Washington, D.C.: U.S. Government Printing Office. 2006.
4. National Center for Health Statistics. Healthy People 2010 Final Review. Hyattsville, MD. 2011.
5. Office of Management and Budget. Standards for maintaining, collecting, and presenting Federal data on race and ethnicity. Federal Register 62: 58781–790. 1997.
6. Office of Management and Budget. Directive No. 15. Race and Ethnic Standards for Federal Statistics and Administrative Reporting. Statistical Policy Handbook. Washington, D.C.: Office of Management and Budget. 1978.
7. Humes KR, Jones NA, Ramirez RR. Overview of race and Hispanic origin: 2010. 2010 Census Briefs. U.S. Census Bureau. March 2011.
8. Rosenberg HM, Maurer JD, Sorlie PD, et al. Quality of death rates by race and Hispanic origin: A Summary of current research, 1999. National Center for Health Statistics. Vital Health Stat 2(128), Hyattsville, MD. 1999. Available from: http://www.cdc.gov/nchs/data/series/sr_02/sr02_128.pdf.
9. National Center for Health Statistics. Vital Statistics of the United States, 1992, Vol. II Mortality, Part A, Technical Appendix. Hyattsville, MD: National Center for Health Statistics. 1996. Available from: http://www.cdc.gov/nchs/data/vsus/mort92_2a.pdf.
10. Sorlie PD, Rogot E, Johnson NJ. Validity of demographic characteristics on the death certificate. Epidemiol 3(2):181–84. 1992.
11. Hogan H. The 1990 Post-enumeration survey: Operations and results. J Am Stat Assoc 88(423):1047–60. 1993.
12. Arias E, Schauman WS, Eschbach K, et al. The validity of race and Hispanic origin reporting on death certificates in the Unites States. National Center for Health Statistics. Vital Health Stat 2(148). Hyattsville, MD. 2008. Available from: http://www.cdc.gov/nchs/data/series/sr_02/sr02_148.pdf.
13. MacDorman, MF, Atkinson, JO. 1998. Infant mortality statistics from the 1996 period linked birth/infant death data Set. Monthly Vital Statistics Report 46(12):1–24. Available from: http://www.cdc.gov/nchs/data/mvsr/supp/mv46_12s.pdf.
14. Kozak LJ. Underreporting of race in the National Hospital Discharge Survey. National Center for Health Statistics. Vital Health Stat Advance Data 265. 1995. Available from: www.cdc.gov/nchs/data/ad/ad265.pdf.
15. National Center for Health Statistics. National Hospital Discharge Survey. 2007. Unpublished tabulation.
16. McCaig LF, Nawar EW. National Hospital Ambulatory Medical Care Survey: 2004 Emergency Department Summary. National Center for Health Statistics. Vital Health Stat Advance Data 372. 2007. Available from: www.cdc.gov/nchs/data/ad/ad372.pdf.
17. Middleton KR, Hing E. National Hospital Ambulatory Medical Care Survey: 2004 Outpatient Department Summary. National Center for Health Statistics. Vital Health Stat Advance Data 373. 2006. Available from: www.cdc.gov/nchs/data/ad/ad373.pdf.

18. Hing E, Cherry DK, Woodwell DA. National Ambulatory Medical Care Survey: 2004 Summary. National Center for Health Statistics. Vital Health Stat Advance Data 374. 2006. Available from: www.cdc.gov/nchs/data/ad/ad374.pdf.
19. Niska R, Bhuiya F, Xu J. National Hospital Ambulatory Medical Care Survey: 2007 Emergency Department Summary. Natl Health Stat Rep 26. Hyattsville, MD: National Center for Health Statistics. 2010. Available from: www.cdc.gov/nchs/data/nhsr/nhsr026.pdf.
20. Hsiao CJ, Cherry DK, Beatty PC, Rechsteiner EA. National Ambulatory Medical Care Survey: 2007 Summary. Natl Health Stat Rep 27. Hyattsville, MD: National Center for Health Statistics. 2010. Available from: www.cdc.gov/nchs/data/nhsr/nhsr027.pdf.
21. Hing E, Hall MJ, Ashman JJ, Xu J. National Hospital Ambulatory Medical Care Survey: 2007 Outpatient Department Summary. Natl Health Stat Rep 28. Hyattsville, MD: National Center for Health Statistics: 2010. Available from: www.cdc.gov/nchs/data/nhsr/nhsr028.pdf.
22. U.S. Department of Health and Human Services, Public Health Service. Healthy People 2000: National Health Promotion and Disease Prevention Objectives. Washington, D.C.: U.S. Government Printing Office. 1991.
23. National Center for Health Statistics. Health United States, 1998, With Socioeconomic Status and Health Chartbook. Hyattsville, MD: National Center for Health Statistics. 1998.
24. NHIS documentation online, available from: www.cdc.gov/NCHS/data/nhis/tecdoc3.pdf.
25. Aud S, Hussar W, Kena G, Bianco K, Frohlich L, Kemp J, Tahan K. The Condition of Education 2011. (NCES 2011-033.) U.S. Department of Education, National Center for Education Statistics . Washington, D.C.: U.S. Government Printing Office. 2011.
26. Sorlie PD, Johnson NJ. Validity of education information on the death certificate. Epidemiol 7(4):437–39. 1996.
27. National Center for Health Statistics. 2003 Revisions of the U.S. Standard Certificates of Live Birth and Death and the Fetal Death Report. 2003. Available from: http://www.cdc.gov/nchs/nvss/vital_certificate_revisions.htm.
28. National Center for Health Statistics. Report of the panel to evaluate the U.S. Standard Certificates and Reports. National Center for Health Statistics. 2000. Available from: http://www.cdc.gov/nchs/data/dvs/panelreport_acc.pdf.
29. Hamilton BE, Martin JA, Ventura SJ. Births: Preliminary data for 2007. Natl Vital Stat Rep 57(12). Hyattsville, MD: National Center for Health Statistics. 2009. Available from: http://www.cdc.gov/nchs/data/nvsr/nvsr57/nvsr57_12.pdf.
30. Xu JQ, Kochanek KD, Murphy SL, Tejada-Vera B. Deaths: Final data for 2007. Natl Vital Stat Rep 58(19). Hyattsville, MD: National Center for Health Statistics. 2010. Available from: http://www.cdc.gov/nchs/data/nvsr/nvsr58/nvsr58_19.pdf.
31. Current and Historical Definitions of Metropolitan and Micropolitan Statistical Areas: http://www.census.gov/population/metro/. (Accessed 09/30/2011.)
32. World Health Organization. International Classification of Impairments, Disabilities, and Handicaps: A Manual of Classification Relating to the Consequences of Disease. Geneva, Switzerland: World Health Organization. 1980.
33. World Health Organization. ICF: International Classification of Functioning, Disability, and Health. Geneva, Switzerland: World Health Organization. 2001.
34. Benson V, Marano MA. Current estimates from the National Health Interview Survey, 1995. National Center for Health Statistics. Vital Health Stat 10(199). 1998. Available from: www.cdc.gov/nchs/data/series/sr_10/10_199_1.pdf.

7. Age Adjustment

Because many Healthy People 2010 objectives have outcomes that vary by age, data for a number of objectives are age adjusted. Age adjustment is a technique to control for differences among populations or changes over time due only to differences in age composition. *Healthy People 2010* uses age-adjusted rates computed by the direct method, which consists in applying the age-specific rates in a population of interest to a standardized age distribution in order to eliminate differences in observed rates that result from age differences in the population composition [1–3].

Age-adjusted rates are useful for comparing two or more populations (such as race/ethnic groups) at a point in time or a single population at two or more points in time. They should not be used to measure absolute magnitude. [Absolute magnitude is best measured by the number of events or by crude (unadjusted) rates.] The actual numerical value of an age-adjusted rate is dependent upon the standard population used and should be viewed as a construct or index rather than a direct or actual measure. It is also important to note that age-adjusted rates may only be compared to rates adjusted to the *same* standard population [4].

In *Healthy People 2010*, many of the mortality objectives are age adjusted, as are many of the objectives that measure health outcomes and risk factors [2,3]. Age-adjusted data may be shown for objectives that target either the total population or a subgroup of the population with a large age range. Objectives or population subgroups that target groups with relatively small age ranges (generally less than 40 years) are not adjusted [4]. Data for older adult age groups (e.g., 50 and over, 65 and over, etc.) are generally age adjusted.

For some population groups, the age-adjusted rates are considerably different from crude rates. This happens because the age distribution of the group is quite different from the age distribution of the standard population. For example, the Hispanic population has a much younger age distribution than the standard population. Consequently, the age-adjusted rates in this population for those outcomes and behaviors that are generally more frequent among the older population are considerably higher than the corresponding crude rates.

With the exception of two objectives (4-1 and 4-7), all age-adjusted rates in *Healthy People 2010* are based on the year 2000 standard population, which was derived from the United States projected 2000 population [2,3]. The selection of the standard age distribution, or standard population, is to some extent arbitrary [5,6]. A number of different standards have been used over the years by Federal and State statistical agencies. Prior to 1999, the National Vital Statistics System used a standard based on the 1940 population, while other agencies of the U.S. Department of Health and Human Services (DHHS) used different standards. Since data year 1999, all DHHS agencies have been using the 2000 standard [7,8].

Those objectives tracked with age-adjusted data are noted in *Tracking Healthy People, Part B*. Data not specifically denoted as age adjusted, should be considered crude (unadjusted) data.

Direct Age Adjustment

The age-adjusted rate (AAR) is a weighted average of age-specific rates where the age-specific weights represent the relative age distribution of a standard population. The AAR is calculated by the direct method using the following formula [9]:

$$AAR = \sum w_{si} \times R_i \ ,$$

where R_i is the age-specific rate for age interval i and w_{si} denotes the standard weight, such that

$$w_{si} = \frac{P_{si}}{\sum P_{si}} \ ,$$

P_{si} denotes population in age interval i in the standard population, $0<w_{si}<1$, and the w_{si} sum to one.

Table 3 shows the 2000 U.S. standard population by single years of age used to derive the age specific weights. In the *Healthy People 2010 Midcourse Review,* the age-adjustment weights for objectives tracked by mortality and health survey data were created from populations rounded to the thousands and aggregated into larger age groups [2]. In early 2006, the method for creating weights was changed [3]. Since that time, whole number populations by single years of age have been used to create the age adjustment weights for age-adjusted Healthy People 2010 objectives. This allows users to age adjust data using any combination of age groups. There is very little difference between the weights using populations rounded to the thousands and weights based on whole number populations. Additional information is available from [10].

Table 3. U.S. Standard Population by Single Years of Age

Age in Years	**2000 U.S. Standard Population**	
	Single Ages to 99	**Single Ages to 84**
00	3,794,901	3,794,901
01	3,758,562	3,758,562
02	3,773,025	3,773,025
03	3,791,001	3,791,001
04	3,869,031	3,869,031
05	3,896,081	3,896,081
06	3,917,855	3,917,855
07	3,978,143	3,978,143
08	3,903,983	3,903,983
09	4,223,778	4,223,778
10	4,230,322	4,230,322
11	4,027,959	4,027,959
12	3,941,299	3,941,299
13	3,923,270	3,923,270
14	3,933,929	3,933,929
15	3,952,423	3,952,423
16	3,853,629	3,853,629

Age in Years	2000 U.S. Standard Population	
	Single Ages to 99	Single Ages to 84
17	4,012,263	4,012,263
18	3,936,904	3,936,904
19	4,064,299	4,064,299
20	4,037,599	4,037,599
21	3,764,802	3,764,802
22	3,555,718	3,555,718
23	3,489,233	3,489,233
24	3,409,873	3,409,873
25	3,421,099	3,421,099
26	3,328,203	3,328,203
27	3,434,987	3,434,987
28	3,450,602	3,450,602
29	4,087,176	4,087,176
30	3,999,004	3,999,004
31	3,810,183	3,810,183
32	3,774,385	3,774,385
33	3,840,938	3,840,938
34	4,086,860	4,086,860
35	4,288,078	4,288,078
36	4,349,620	4,349,620
37	4,469,476	4,469,476
38	4,290,207	4,290,207
39	4,782,575	4,782,575
40	4,666,685	4,666,685
41	4,493,582	4,493,582
42	4,487,560	4,487,560
43	4,424,004	4,424,004
44	4,407,398	4,407,398
45	4,268,017	4,268,017
46	4,033,859	4,033,859
47	3,958,468	3,958,468
48	3,681,489	3,681,489
49	3,863,960	3,863,960
50	3,720,935	3,720,935
51	3,504,329	3,504,329
52	3,475,657	3,475,657
53	3,754,218	3,754,218
54	2,769,220	2,769,220
55	2,749,739	2,749,739
56	2,786,795	2,786,795
57	2,947,472	2,947,472
58	2,404,462	2,404,462

Age in Years	2000 U.S. Standard Population	
	Single Ages to 99	Single Ages to 84
59	2,418,766	2,418,766
60	2,259,141	2,259,141
61	2,179,759	2,179,759
62	2,132,873	2,132,873
63	2,030,730	2,030,730
64	2,051,769	2,051,769
65	2,033,933	2,033,933
66	1,862,107	1,862,107
67	1,849,893	1,849,893
68	1,788,769	1,788,769
69	1,875,238	1,875,238
70	1,843,087	1,843,087
71	1,784,744	1,784,744
72	1,802,080	1,802,080
73	1,674,285	1,674,285
74	1,621,378	1,621,378
75	1,610,943	1,610,943
76	1,530,137	1,530,137
77	1,450,062	1,450,062
78	1,456,186	1,456,186
79	1,367,231	1,367,231
80	1,172,978	1,172,978
81	1,065,672	1,065,672
82	963,587	963,587
83	890,893	890,893
84	807,104	807,104
85	693,158	4,259,173
86	607,940	
87	536,762	
88	452,814	
89	387,893	
90	327,827	
91	273,709	
92	226,917	
93	180,330	
94	143,772	
95	118,131	
96	88,924	
97	65,909	
98	46,278	
99	37,194	
100+	71,615	

Age in Years	2000 U.S. Standard Population	
	Single Ages to 99	Single Ages to 84
Total	274,633,642	274,633,642

Source: U.S. Census Bureau, Population Estimates Program.

Mortality

There are about 55 Healthy People 2010 objectives that monitor mortality outcomes [2,3]. Most of these use data from the National Vital Statistics System, of which 27 are tracked with age-adjusted death rates (see *Tracking Health People, Appendix D*). The remaining mortality objectives are measured using either:

- Numbers of deaths
- Age-specific death rates
- Maternal/infant mortality rates, for which births are the denominator, or
- Crude death rates from data systems other than the National Vital Statistics System, such as the Fatality Analysis Reporting System or the Census of Fatal Occupational Injuries.

The measurement details for each objective are specified in the operational definitions [1].

After publication of *Healthy People 2010* [1], all mortality baselines were revised to data year 1999 to accommodate the change to ICD-10 (see Section 8, below). Age-adjusted death rates used for Healthy People 2010 are based on the 2000 population standard. Therefore, they differ from rates shown in previous Healthy People 2000 reports [11], which were based on the 1940 standard population. Thus, the rates computed for Healthy People 2000 cannot be used in trend comparisons with rates computed for Healthy People 2010.

Health Surveys

A number of other Healthy People 2010 objectives use aged-adjusted data from national health surveys such as the National Health Interview Survey, the National Hospital Discharge Survey, the National Hospital Ambulatory Medical Care Survey, the Behavioral Risk Factor Surveillance System (BRFSS), and the National Health and Nutrition Examination Survey. Data for these objectives are age adjusted to the 2000 standard population, using the above equations with R_i as the age-specific rate for the health status, health behavior, or health care utilization variable, as appropriate.

The age groups used to adjust Healthy People 2010 survey data are shown in the operational definition for each objective. The specific grouping used depends on the data system and the population targeted by the objective [11]. While the same standard population is employed, the age groups used to adjust survey data may differ from those used to adjust mortality data. In general, to maximize the stability of the rates, fewer age groups are used. Differences in adjusted rates resulting from the different specific age groups used should be relatively small.

In some cases, the applicable age range for the objective may not be the total population. For example, an objective may refer to persons aged 18 years and over, females aged 40 years and over, or persons aged 45–74 years. In these cases, the weights are based on the population in the specified age range selected for adjustment.

When the denominator for an objective is persons with a chronic disease (diabetes, cancer, chronic kidney disease, arthritis, coronary heart disease, and hypertension), the data for age groups under 45 years (or, in some cases, 40 years) are aggregated into a single group for the age-adjustment calculation. This is done to stabilize the age-adjusted rates. The age distribution of persons with chronic diseases tends to differ considerably from the standard population used for age adjustment. Using the standard age groups to age-adjust an objective with a chronic disease denominator places relatively large weights on the younger age groups. The relatively small numbers of people with those conditions in these age groups may result in highly variable rates. Combining the younger age groups increases stability and reliability.

More information on age adjustment of survey data for Healthy People 2010 can be found in [11].

References

1. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.
2. U.S. Department of Health and Human Services. Healthy People 2010 Midcourse Review. Washington, D.C.: U.S. Government Printing Office. 2006.
3. National Center for Health Statistics. Healthy People 2010 Final Review. Hyattsville, MD. 2011.
4. Shryock HS, Siegel JS. The Methods and Materials of Demography. San Diego, CA: Academic Press. 1976.
5. Feinleib M, Zarate AO (eds). Reconsidering age adjustment procedures: Workshop proceedings. National Center for Health Statistics. Vital Health Stat 4(29). 1992. Available from: www.cdc.gov/nchs/data/series/sr_04/sr04_029.pdf.
6. Curtin LR, Maurer J, Rosenberg HM. On the selection of alternative standards for the age-adjusted death rate. Am Stat Assoc, Proc Soc Stat Sect: 218–23. 1980.
7. Anderson RN, Rosenberg HM. Report of the second workshop on age adjustment. National Center for Health Statistics. Vital Health Stat 4(30). 1998. Available from: http://www.cdc.gov/nchs/data/series/sr_04/sr04_030.pdf.
8. Anderson RN, Rosenberg HM. Age standardization of death rates: Implementation of the year 2000 standard. Natl Vital Stat Rep 47(3). Hyattsville, MD: National Center for Health Statistics. 1998. Available from: http://www.cdc.gov/nchs/data/nvsr/nvsr47/nvs47_03.pdf.
9. Curtin LR, Klein RJ. Direct standardization (age-adjusted death rates). Healthy People Statistical Notes, no. 6, revised. Hyattsville, MD: National Center for Health Statistics. 1995. Available from: http://www.cdc.gov/nchs/data/statnt/statnt06rv.pdf.
10. Surveillance, Epidemiology, and End Results (SEER). Standard Populations (Millions) for Age Adjustment. 2006. Available from: http://seer.cancer.gov/stdpopulations/.
11. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People Statistical Notes, no. 20. Hyattsville, MD: National Center for Health Statistics. 2001. Available from: http://www.cdc.gov/nchs/data/statnt/statnt20.pdf.

8. Mortality and Morbidity Classification

Mortality Data

Data for objectives that monitor specific causes of death are classified and coded according to the World Health Organization's (WHO) Tenth Revision of the International Classification of Diseases (ICD-10) [1]. The ICD is a classification system that provides basic guidance for coding and classifying causes of death. It includes disease, injury, and poisoning categories, as well as the rules used to select the single underlying cause of death from the several diagnoses that may be reported on the death certificate. The ICD also includes definitions, tabulation lists indicating cause-of-death groupings used to present mortality data, and the format of the medical certificate of death. Use of the ICD for the classification and coding of mortality statistics is required under an agreement between the United States and the WHO [2].

Since 1900, the ICD for mortality has been revised approximately every 10 years, with the exception of the 20-year interval between the Ninth and Tenth revisions. The revisions are essential to ensure that disease classifications are consistent with advances in medical science and changes in diagnostic practice. ICD-10 was implemented in the United States effective with deaths occurring in 1999 [3].

The original Healthy People 2010 baselines for mortality objectives were based on 1997 and 1998 data coded to the Ninth Revision of the ICD (ICD-9) [4,5]. The introduction of a new ICD revision creates discontinuities in time series trends for causes of death because of the reclassification and changes in coding rules. Consequently, to minimize these discontinuities, the baseline data year for cause-specific mortality objectives was revised to 1999. Subsequent tracking data for these objectives are classified according to ICD-10. The specific ICD-10 codes used for each mortality objective are shown in *Tracking Healthy People, Part B* and *Tracking Healthy People, Appendix C*.

Data for most Healthy People 2010 mortality objectives are based on the underlying cause of death. The underlying cause of death is defined as the disease or injury that initiated the sequence of events leading directly to death or as the circumstances of the violence or accident that produced the fatal injury [1]. It is selected from the conditions entered by the physician in the cause of death section on the death certificate. When more than one cause is entered by the physician, the underlying cause is determined by the sequence of conditions on the certificate, provisions of the ICD, and associated selection rules and modifications [3]. Generally, more information is reported on the death certificate than is directly reflected in the underlying cause of death. This is captured in the multiple cause-of-death data. Several objectives use all mentions of a cause (or "multiple" cause) on the death certificate. Specific objectives tracked by multiple-cause statistics are noted in the operational definitions.

The United States Standard Certificate of Death was revised in 2003 to improve the quality of data reported on the death certificate and to facilitate the collection of data needed to address coding changes resulting from the implementation of ICD-10. Some of the revisions include: the provision of additional space in the cause of death section to give the attending physicians more opportunity to list contributing conditions; the addition of a question regarding tobacco use to help reduce underreporting of tobacco use as a contributing factor to death; the addition of a question regarding pregnancy status of female decedents to improve the reporting of maternal deaths; and the addition of a question to gather information about the decedent's role in transportation accidents. The education item was revised from an open-ended item about years of education to check boxes of educational degree categories, consistent with U.S. Census Bureau data. In addition, the race and Hispanic origin questions were changed from open-ended items to check boxes in order to replicate the format of the race and Hispanic origin items on the Census questionnaire [6]. The transition to the 2003 Standard Certificate of Death by the states is occurring over multiple years. Twenty-two states and the District of Columbia have implemented the revision in 2007 [7]. Beginning in 2003, education data for mortality objectives from the National Vital Statistics System were suppressed. Educational attainment data collected using the 2003 version are not comparable with data collected using the 1989 version [7].

Morbidity Data

Baseline data for cause-specific morbidity objectives are coded to International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) [8]. The specific ICD-9-CM codes used are shown in the operational definition for each morbidity objective and in *Tracking Healthy People, Appendix D*.

ICD-9-CM is a clinical modification of ICD-9. The ICD-9-CM coding system includes a fifth digit, thus providing greater specificity and detail than ICD-9. ICD-9-CM is intended to serve as a tool in the classification of morbidity data for indexing of hospital medical records, medical care review, and ambulatory and other medical care programs, as well as for basic health statistics. It is used to code and classify morbidity data from inpatient and outpatient records, physicians’ offices, long term care facilities and most health surveys [8].

Code assignment using ICD-9-CM is based on official national coding guidelines. The guidelines for selecting the “first-listed” or principal diagnosis for morbidity records differ from those used to select the underlying cause of death on death records. Under morbidity coding rules, the first listed or principal diagnosis is that condition established after study to be chiefly responsible for occasioning the admission to the hospital or the encounter with the health care provider for care. In some instances the principal diagnosis may be a manifestation of the disease rather than the underlying cause. For example, if a patient with a primary malignant neoplasm with metastasis is admitted to receive treatment directed toward a secondary site, the secondary site would be designated on the hospital discharge form as the principal diagnosis [9].

In general, the Healthy People 2010 morbidity objectives are tracked using the principal diagnosis. However, in some cases "all-listed" diagnoses, which include the principal and all other diagnoses appearing on the medical record, are used (as many as seven to 10 diagnoses may appear in some records). Specific objectives tracked by all-listed diagnoses are noted in the operational definitions.

Additional codes have been added and code changes have been made to the ICD-9-CM since its implementation in 1986. A conversion table for diagnosis and procedure code changes between 1986 and the current data year is available to assist users in data retrieval. The table shows the date the new code became effective and its previously assigned code equivalent. The latest additions to the classification appear in bold print [9].

A clinical modification of ICD-10 (ICD-10-CM) has been developed as a replacement for ICD-9-CM. The draft version of the ICD-10-CM was open to public comment and underwent field testing by the American Hospital Association and the American Health Information Management Association. Additional modifications have been made to the ICD-10-CM based on comments and suggestions from the comment period and field tests as well as new concepts based on the established update process for ICD-9-CM and ICD-10. The 2011 release of the ICD-10-CM is now available, however, it has not yet been implemented in the United States. On January 16, 2009, HHS published a final rule adopting ICD-10-CM to replace ICD-9-CM in HIPAA transactions, effective implementation date of October 1, 2013 [10].

Once implemented, revised coding guidelines, training materials and crosswalks between ICD-9-CM and ICD-10-CM will be made available on the NCHS website.

Data for the Healthy People 2000 morbidity objectives from the sources listed above were also coded according to ICD-9-CM, although in a few cases, the specific codes used for the Healthy People 2010 objectives are different from those used for the comparable Healthy People 2000 objective. These differences are noted in the operational definition for the objective in *Tracking Healthy People, Part B*.

References

1. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision. Geneva, Switzerland: World Health Organization. 1992.
2. Kochanek KD, Murphy SL, Anderson RN, Scott C. Deaths: Final data for 2002. Natl Vital Stat Rep 53(5). Hyattsville, MD: National Center for Health Statistics. 2004. Available from: http://www.cdc.gov/nchs/data/nvsr/nvsr53/nvsr53_05acc.pdf.
3. Hoyert DL, Arias E, Smith BL, et al. Deaths: Final Data for 1999. Natl Vital Stat Rep 49(8). Hyattsville, MD: National Center for Health Statistics. 2001. Available from: http://www.cdc.gov/nchs/data/nvsr/nvsr49/nvsr49_08.pdf.
4. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.
5. World Health Organization. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, Based on the Recommendations of the Ninth Revision Conference. Geneva, Switzerland: World Health Organization. 1977.
6. National Center for Health Statistics. Report of the Panel to Evaluate the U.S. Standard Certificates. Hyattsville, MD: National Center for Health Statistics. 2001. Available from: www.cdc.gov/nchs/data/dvs/panelreport_acc.pdf.
7. Xu JQ, Kochanek KD, Murphy SL, Tejada-Vera B. Deaths: Final data for 2007. Natl Vital Stat Rep 58(19). Hyattsville, MD: National Center for Health Statistics. 2010. Available from: http://www.cdc.gov/nchs/data/nvsr/nvsr58/nvsr58_19.pdf.
8. U.S. Department of Health and Human Services. International Classification of Diseases, Ninth Revision, Clinical Modifications. 6th ed. DHHS publication no. 96-1260. Washington, D.C.: Health Care Financing Administration. 1997.
9. National Center for Health Statistics. Conversion Table of New ICD-9-CM Codes. 2011. Available from: http://www.cdc.gov/nchs/data/icd9/CNVTB12.pdf.
10. National Center for Health Statistics. International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM). Hyattsville, MD: National Center for Health Statistics. 2011. Available from: http://www.cdc.gov/nchs/icd/icd10cm.htm.

9. National Data

Data used to track the Healthy People 2010 objectives are based on events occurring in the 50 states and the District of Columbia, where available. Unless specifically noted, all objectives exclude data for U.S. territories. The data used to track most population-based Healthy People 2010 objectives are derived from either a national census of events (for example, the National Vital Statistics System and the National Notifiable Disease Surveillance System) or from nationally representative sample surveys (for example, the National Health Interview Survey and the National Survey on Drug Use and Health).

For some objectives, however, complete national data are not available and data for selected states and/or areas are used to monitor the objectives. In these cases, the coverage area is described with the data for the objective and in the operational definitions. Examples of these data systems include the Adult Spectrum of Disease Project and the Health Care Cost and Utilization Project. Data for these objectives may not be representative of the United States as a whole. If the data used to track an objective are not nationally representative, the number of states in the reporting area is noted. If during the decade national data become available, they will be used to track the objectives.

For some national data systems that cover the entire United States, such as the National Vital Statistics System and the Behavioral Risk Factor Surveillance System, data are not available for some variables for all states. This is either because data for a specific variable are not collected by some states or because the quality of data for some states is not sufficient to produce reliable estimates for some variables. For example, beginning in 2003, education data for mortality objectives from the National Vital Statistics System have been suppressed. The educational attainment item was changed in the new U.S. Standard Certificate of Death in 2003 to be consistent with the U.S. Census Bureau data and to improve the ability to identify specific types of educational degrees. Many states, however, are still using the 1989 version of the U.S. Standard Certificate of Death, which focuses on highest school grade completed. As a result, educational attainment data collected using the 2003 version are not comparable with data collected using the 1989 version. As a second example regarding the Behavioral Risk Factor Surveillance System, diabetes variables were only available for 37 states and the District of Columbia in 2008 which is the most recent data year available for Healthy People 2010. The number of reporting states can vary from year to year. This information is also shown in the operational definitions for selected objectives.

10. State and Local Data

Healthy People has provided a framework for national, State, and local health agencies and non-governmental organizations to assess health status, health behaviors, and services and to plan and evaluate health promotion programs [1,2]. The national Healthy People initiatives have served as a "menu" for identifying state and local priorities and selecting objectives that are most relevant to specific states, communities, and specific settings, e.g., schools, worksites, etc., and health care delivery systems [3,4]. Twenty-two percent of the objectives included in the Department of Health and Human Services' strategic plan, developed in accordance with the Government Performance and Results Act of 1993, were adopted from Healthy People 2000.

This focus on performance has prompted State and local health agencies to shift from their emphasis of primarily providing services to one that conducts needs assessment and quality assurance [5,6]. Health care delivery organizations have also experienced this shift and have increased efforts to collect standardized data on patients, services, and outcomes. The shift in emphasis on data collection and analysis for purposes of assessment and evaluation has increased the need to address the associated issues of data availability, validity, reliability, comparability, and utilization.

The large number and diversity of State and local health agency structures and resources amplify the impact of these data issues when comparing Healthy People objectives, and progress among states or between a state and the nation. In order to address the need to develop capacity for tracking at the state and local level, Healthy People 2010 includes a separate Focus Area, Focus Area 23, Public Health Infrastructure, aimed at improving infrastructure and surveillance capability [6,7].

Some key areas where these issues need to be examined at state and local levels are discussed below.

Objective Wording/Operational Definition

Many State and local agencies and organizations have tailored the Healthy People 2010 objectives to better focus on specific concerns of their constituents. These modifications may more effectively address the health concerns of the state or local population, but they also reduce comparability when evaluating objective progress relative to the nation, other states, or localities. Examples include objective 8-18 which tracks the proportion of persons living in homes tested for radon at the national level. Some states have adopted this objective verbatim, while others have included schools or day care centers within the same objective. In addition, the operational definitions and data sources for the same or similar national and state/local objectives may vary considerably. Self-reported data from a household survey is used to monitor the national radon objective, whereas some states use data based on actual installation of radon monitors and picocurie information collected.

Population Data/Race and Ethnicity Reporting

Many Healthy People objectives are population based and are expressed as rates, e.g., objective 3-2, lung cancer deaths per 100,000, where the denominator is a population estimate. National, State, and local health agencies primarily rely on population estimates produced by the United States Census Bureau. The Bureau provides population counts from the decennial census for the nation, states, counties, and large municipalities. It also produces annual postcensal estimates for the years following the census (see Section 5: Population Estimates). These estimates are available by sex, age, race and ethnicity (see Section 6: Population Template). However, the sizes of some racial groups, e.g., American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander are relatively small, even at the national level, and are distributed unevenly across state and local areas. This precludes many jurisdictions from producing reliable rates for objectives that focus on these populations.

"Rare" Events/Confidentiality

Some Healthy People objectives, e.g., objective 13-14 HIV deaths, address important, sensitive health issues which are relatively rare events. Small numbers of HIV deaths in a county or municipality with a relatively small population may result in unreliable, non-representative rates. Reporting these rates by certain characteristics or geocoding and displaying maps of the distribution of sensitive or rare events may jeopardize privacy of affected individuals or confidentiality of the source data. It is often necessary to aggregate data over geographic areas, demographic characteristics, and/or data years to address these issues.

Age Adjustment

In general, states and localities age-adjust mortality data to the same standard population used for the national data (see Section 7: Age Adjustment).

Data Sources

The availability and comparability of data for national, state, and local monitoring of Healthy People objectives vary considerably. Some data, especially vital statistics, are readily available at all geographic levels. The standardization of vital statistics data contributes to their comparability across jurisdictions.

Many Healthy People objectives are monitored using data from the National Health Interview Survey (NHIS). Some of these objectives are also monitored at state and local levels using data from the Behavioral Risk Factor Surveillance System (BRFSS). In general, it should be noted that differences in the data collection methods, i.e., household interview vs. telephone interview, and wording of questions used to monitor the same objectives at the national vs. the state level can affect the comparability of the information collected. In addition, some objectives monitored with identical questions in the NHIS and the BRFSS, e.g., objective 15-4, firearm storage in homes, are only included periodically in a specific rotating module of the BRFSS or supplements to the NHIS.

Not all states use these modules and/or the year of the "rotation" may not coincide with national data from the NHIS. This limits comparability between national and state data.

Other Healthy People objectives are monitored using composite data sources.The national data from these systems are aggregated from data collected at state or local levels. Unlike the vital statistics data which include all births and deaths, several of these systems are samples of events that use different data collection and analysis methods among states or communities. For example, the National Water Quality Inventory data are compiled from state data on "assessed" rivers, lakes, and estuaries. States vary in the proportions and the specific bodies of water they assess across time. Hence, both state and national estimates may be subject to considerable variation. This affects the quality and comparability of national, state, and local data.

Many Healthy People 2000 objectives lacked data at the state and local jurisdiction level to monitor progress toward some of the Healthy People 2000 objectives. This prompted the development of Healthy People 2000 Priority Data Needs, which identified sources of state and local data that could be used to track important health issues, such as adult immunization and access to primary health care [8]. While local data are not yet available, state data are currently available through DATA2010 for selected objectives (see Section 12: Healthy People 2010 Database). State data can be accessed through all table generating options available in DATA2010.

References

1. American Public Health Association. Healthy Communities: Guidelines for Community Attainment of the Year 2000 National Health Objectives. Washington, D.C.: American Public Health Association. 1991.
2. U.S. Department of Health, Education, and Welfare, Public Health Service. Promoting Health/Preventing Disease: Objectives for the Nation. Washington, D.C.: U.S. Government Printing Office. 1980.
3. U.S. Department of Health and Human Services, Public Health Service. Healthy People 2000: National Health Promotion and Disease Prevention Objectives. Washington, D.C.: U.S. Government Printing Office. 1991.
4. Institute of Medicine. The Future of Public Health. Washington, D.C.: National Academy Press. 1988.
5. National Center for Health Statistics. The CDC Assessment Initiative: A Summary of State activities. Statistics and Surveillance Notes, no. 7. Hyattsville, MD: National Center for Health Statistics. 1995. Available from: http://www.cdc.gov/nchs/data/stsurv/stsurv7.pdf.
6. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.
7. Freedman MA. Health status indicators for the year 2000. Healthy People Statistical Notes, no. 1. Hyattsville, MD: National Center for Health Statistics. 1991. Available from: http://www.cdc.gov/nchs/data/statnt/statnt01acc.pdf.
8. Kim I, Keppel GK. Priority data needs: Sources of national, state, and local-level data and data collection systems. Healthy People Statistical Notes, no. 15. Hyattsville, MD: National Center for Health Statistics. 1997. Available from: http://www.cdc.gov/nchs/data/statnt/statnt15.pdf.

11. Variability of Estimates and Data Suppression

Two main types of data systems are used to track Healthy People 2010 objectives: *sample surveys* and *population-count systems* (also called surveillance systems). Evaluating data from both types of systems requires consideration of variability. For sample surveys, sampling error (also called sampling variability) is of interest. For population-count systems, random variation is of interest. Issues of data quality (for example, item non-response, bias, non-representativeness) can affect data from both types of sources.

Sample Surveys

For many health outcomes, assessing all individuals in a population may be impossible, impractical, expensive, or inaccurate. Therefore, it is usually advantageous to study a sample of the original population. Much of the data used to monitor the Healthy People 2010 objectives are derived from sample surveys (for example, the National Health Interview Survey and the National Survey on Drug Use and Health) that make estimates for a population from a representative sample of respondents. These estimates are subject to sampling error. One commonly-used measure of sampling error is the standard error. The standard error represents the variation in an estimate that can occur by chance since only a sample of the population is surveyed rather than the entire population. Assuming a normal distribution of events, the chances are about 68 in 100 that an estimate from the sample would differ from a complete census by less than the standard error. The chances are about 95 in 100 that the difference would be slightly less than twice the standard error. This is often referred to as the 95 percent confidence interval, where the estimate is expressed as a range of the observed rates, approximately ±1.96 standard errors.

To properly interpret differences between rates for different population groups or changes over time in data derived from sample surveys, it is important to consider the variation associated with each rate. *Healthy People 2010* [1] uses a population template that includes detailed racial, ethnic and socioeconomic categories for all population-based objectives; see Section 6: Population Template. This template necessitates the tabulation of data for relatively small population groups. These data are often associated with large standard errors. Thus, apparent differences between population groups or between a population group in the template and the total population may be within expected sampling error. Standard errors should be considered when evaluating progress or comparing population groups for objectives using survey data. For ease of presentation, the standard errors associated with the estimates for the Healthy People 2010 objectives do not appear in either *Healthy People 2010* [1] or *Tracking Healthy People 2010*. However, where available, they are included in the Healthy People 2010 database, DATA2010; see Section 12: Healthy People 2010 Database. More information on the sample design and variance estimation for some of the major data systems used to monitor the Healthy People 2010 objectives can be found in *Tracking Healthy People, Part C*, and in other publications [2–9].

Population-Count Systems

Some of the data systems used to track the Healthy People 2010 objectives are based on complete counts of events occurring to the population (for example, the National Vital Statistics System, the HIV-AIDS Reporting System, and the United States Renal Data System). As such, these data are not subject to sampling error, although they are subject to errors in the registration process. However, when the estimates are used for analytical purposes, such as the comparison of rates over time or among groups or areas, the number of events that actually occurred may be considered as one of a large series of possible results that could have arisen under the same circumstances. This is known as random variation. When the number of events is large, random variation is usually small. However, when the number of events is small (fewer than 100) and the probability of such an event is small, random variation can be substantial and considerable caution must be used in interpreting the change described by the estimates. In these cases, it is desirable to compute the standard error of the rates and use that computation in the comparison of interest. Standard errors for rates derived from population-count systems, where available, are included in DATA2010. More information on random variation and small numbers can be found in the "technical notes" section of the annual National Vital Statistics Reports [10,11].

Data Suppression

Healthy People 2010 shows data for all of the groups included in the population template, when available [1]. However, in some instances data are not available and one of the following symbols is shown in place of a data value:

- 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. There are three main reasons a Healthy People 2010 statistic is suppressed (shown as DSU):

- The number of events is too small to produce a reliable estimate or may violate confidentiality requirements.
- The sample design does not produce representative estimates for a particular group.
- There is a high item non-response rate or a large number of unknown entries.

Different criteria for data suppression have been adopted by the various data systems used to monitor the Healthy People 2010 objectives. Some sample surveys use a single criterion for data suppression. For example, data from the Behavior Risk Factor Surveillance System are considered unreliable and are suppressed if the denominator is based on fewer than 50 sample cases. Other sample surveys use a combination of criteria. For example, data from the National Health Interview Survey are suppressed if the denominator is based on fewer than 50 events or if the relative standard error is greater than 30%.

For population-count systems that are based on a complete census, typically a single criterion for data suppression based on the number of events is used. For example, the National Vital Statistics System considers rates based on fewer than 20 events to be unreliable.

In presenting data for the Healthy People 2010 objectives, NCHS adheres to the specific criteria for data suppression delineated by each data system. More information on data suppression, including a summary of the criteria used by the "major data systems" for Healthy People 2010, has been published elsewhere [12].

For most objectives, an estimate or count based on a single year can be shown. However, for some objectives (or subgroups in the population template) that are based on relatively few events, multiple years of data are used to produce more stable estimates. For example, all of the objectives measured by the National Health and Nutrition Examination Survey use estimates based on either two, four, or six years of data.

Special Situations in Data Suppression

The implementation of the 1997 OMB race standard resulted in changes in how information about race was collected by each data source over the decade. In Tracking Healthy People, the minimum race template for population based objectives has been modified since the beginning of the decade to reflect the 1997 OMB standard (see Section 6: Population Template).

Data in DATA2010 that were collected by a given data source prior to the data source implementation of the 1997 OMB race standard are displayed in the minimum race template, and footnoted to reflect the differences in how these data were collected compared to how data are collected after the implementation of the 1997 OMB standard. The standard was implemented in different years by different data sources; and some data sources, such as the National Vital Statistics System, are still in the process of implementing the changes.

As a result of the change in data collected by different data sources, DATA2010 users will notice different 'data suppression' patterns for data from each data set for before and after the 1997 OMB race standard implementation. The symbols (DNC, DNA, DSU) used in DATA2010 when data are not available are described above.

Special situations are described below:

National Health Interview Survey

The National Health Interview Survey implemented the 1997 OMB race standard in 1999. Prior to the implementation of the new race standard, NHIS had allowed interviewers to record multiple race responses. However, data lines in DATA2010 for the years prior to the implementation of the new standard in NHIS display the data for the categories 'Two or more races,' Black or African American; White, and American Indian and Alaska Native; White, as DNC or data not collected. These data are displayed as DNC because at that time, the Healthy People 2010 template had not yet been updated to include multiple race categories.

National Vital Statistics System – Mortality

The National Vital Statistics System began a partial implementation of the 1997 OMB race standard for 2003. Currently, data collected by the National Vital Statistics System under the 1997 OMB standard are bridged back to be comparable with the data collected prior to the 1997 OMB standard. NVSS data are also displayed in DATA2010 using the race template that was in the original *Healthy People 2010* publication (see Section 6: Population Template) [1]; this is consistent with how data are displayed by this system.

DATA2010 displays data for the National Vital Statistics System—Mortality data for years 1998 forward. There are different population denominators available from the U.S. census for the data displayed for the years before 2000 and for years 2000 and after. Before the year 2000 the population for the Asian and Pacific Islander category was combined, and beginning with the 2000 census this category was split into two categories, Asian, and Native Hawaiian or Other Pacific Islander. In DATA2010, the Asian and Native Hawaiian or Other Pacific Islander categories are displayed as DNC in 1998 and 1999 since denominator information was not available at that time, and they are displayed as DNA in the years 2000 forward. When the National Vital Statistics System – Mortality has fully implemented the 1997 OMB race standard DATA2010 will display data for both the Asian and Native Hawaiian or Pacific Islander categories.

References

1. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. With Understanding and Improving Health and Objectives for Improving Health. 2 vols. Washington, D.C.: U.S. Government Printing Office. November 2000.
2. Fleishman JA. Demographic and clinical variations in health status. MEPS Methodology Report No. 14. AHRQ Pub. No. 05-0022. Agency for Healthcare Research and Quality: Rockville, MD. 2005.
3. McDowell MA, Fryar CD, Hirsch R, Ogden CL. Anthropometric reference data for children and adults: U.S. population, 1999–2000. National Center for Health Statistics. Vital Health Stat Advance Data 361. 2005. Available from: www.cdc.gov/nchs/data/ad/ad361.pdf.
4. Botman SL, Moore TF, Moriarty CL, Parsons VL. Design and estimation for the National Health Interview Survey, 1995–2004. National Center for Health Statistics. Vital Health Stat 2(130). 2000. Available from: www.cdc.gov/nchs/data/series/sr_02/sr02_130.pdf.
5. 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. Available from: www.cdc.gov/nchs/data/series/sr_01/sr01_039.pdf.
6. Smith PJ, Hoaglin DC, Battaglia MP, et al. Statistical methodology of the National Immunization Survey, 1994–2002. National Center for Health Statistics. Vital Health Stat 2(138). 2005. Available from: www.cdc.gov/nchs/data/series/sr_02/sr02_138.pdf.
7. Groves R, Benson G, Mosher WD, et al. Plan and operation of cycle 6 of the National Survey of Family Growth. National Center for Health Statistics. Vital Health Stat 1(42). 2005. Available from: www.cdc.gov/nchs/data/series/sr_01/sr01_042.pdf.
8. Kolbe LJ, Kann L, Collins JL, Small ML, Pateman BC, Warren CW. The School Health Policies and Programs Study (SHPPS): Context, methods, general findings, and future efforts. J Sch Health Suppl 7(71). 2001.

9. Centers for Disease Control and Prevention. Methodology of the Youth Risk Behavior Surveillance System. Morb Mortal Wkly Rep 53(RR12). 2004. Available from: www.cdc.gov/mmwr/pdf/rr/rr5312.pdf.
10. Kochanek KD, Murphy SL, Anderson RN, Scott C. Deaths: Final Data for 2002. Natl Vital Stat Rep 53(5). Hyattsville, MD: National Center for Health Statistics. 2004. Available from: www.cdc.gov/nchs/data/nvsr/nvsr53/nvsr53_05.pdf.
11. Martin JA, Hamilton BE, Sutton PD, et al. Births: Final data for 2003. Natl Vital Stat Rep 54(2). Hyattsville, MD: National Center for Health Statistics. 2005. Available from: http://www.cdc.gov/nchs/data/nvsr/nvsr54/nvsr54_02.pdf.
12. Klein RJ, Proctor SE, Boudreault MA, Turczyn KM. Healthy People 2010 criteria for data suppression. Healthy People Statistical Notes, no. 24. Hyattsville, MD: National Center for Health Statistics. June 2002. Available from: www.cdc.gov/nchs/data/statnt/statnt24.pdf.

12. Healthy People 2010 Database

The Healthy People 2010 database, DATA2010, is an interactive, on-line database developed by the Centers for Disease Control and Prevention, National Center for Health Statistics (NCHS). It can be accessed from the Healthy People 2010 website or from http://wonder.cdc.gov/data2010.

DATA2010 contains the baseline and tracking data for all measurable Healthy People 2010 objectives. National data are available for all of the measurable objectives. State data are available for a subset of the measurable objectives. Socio-demographic data for population-based objectives are also provided; data are shown by race and Hispanic origin, sex, and socio-economic status (education or income). Some objectives have additional data by age, disability status, and residency location (urban/rural). Through a series of menus and drop-down boxes, users can obtain data for:

- All objectives for one of the 28 Focus Areas
- Objectives used to track the ten Healthy People 2010 Leading Health Indicators
- All objectives for a particular data source
- A single objective within a Focus Area
- Objectives containing a specific word or phrase
- All objectives for a specific population group (e.g., adolescents, Hispanics, or women).

The standard data table output includes the following information:

- Objective number
- Objective text
- Baseline year and data
- Tracking data for subsequent years
- 2010 Target
- Footnotes
- Data source(s).

The standard data tables can be constructed in ASCII, comma-delimited, or HTML format for use in common software applications such as EXCEL, Lotus 1-2-3, and SAS. Users can also download statistical data spreadsheets in EXCEL format that contain unrounded data and standard errors (where available). Bar/line charts can be created in DATA2010 using an advanced selection option.

In addition to data, DATA2010 contains other technical information related to tracking the Healthy People 2010 objectives. For instance, users can obtain operational definitions for each objective that provide detailed information including the description of numerator, denominator, population, questions used to collect data, ICD-9 and ICD-10 codes for morbidity and mortality objectives, and the age-adjustment categories used for age-adjusted measures.

DATA2010 was updated on a quarterly basis to provide the most accurate data for Healthy People 2010 objectives. New data and revisions to data previously shown were added during each update.