

Population Group Definitions and Categorizations for Healthy People 2020 Final Review

The Healthy People 2020 (HP2020) Final Review includes an evaluation of final progress status for population groups defined by six characteristics: sex, race and ethnicity, educational attainment, family income, disability status, and geographic location. The population group analyses included a subset ($n = 611$; 55.0%) of the 1,111 measurable objectives in the HP2020 overall Final Progress Table, although the number of population groups included for each objective varied based on data availability. Additionally, the definition of population groups varied across the data sources used in HP2020.

This document describes the definitions and categorization of population groups in the Progress by Population Group analysis in the HP2020 Final Review. It also describes how targets that were developed based on the total population covered by the objectives were used in the population group analyses. Users should also refer to the [General Data Issues](#) document (1), which provides additional detailed information about population groups and the HP2020 data sources.

Sex

Data for males and females were reported separately in HP2020 when available and applicable. Sex is self-reported by respondents in surveys and other interview-based data collection systems. Sex is often recorded by health care professionals based on observation or written records in many surveillance systems. Sex may also be reported by proxies (for example, the adult head of family or household may report the sex of children). In the Progress by Population Group analyses, 516 objectives had data available for males, and 517 objectives had data available for females.

Race and Ethnicity

Federal standards on record keeping and the collection and presentation of data on race and ethnicity were revised by the Office of Management and Budget (OMB) in 1997. Previously, race and ethnicity standards specified four single-race categories, and only one category could be selected: American Indian or Alaska Native, Asian or Pacific Islander, Black, and White. The revised 1997 standards allow respondents to select one or more of the following five categories when reporting their race. These five categories are the minimum set for data on race for federal statistics, program administrative reporting, and civil rights compliance reporting.

American Indian or Alaska Native: A person having origins in any of the original peoples of North America 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.

Collection of additional detail on race or ethnicity is permitted if the additional categories can be combined into the minimum categories. 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.

The standards on Hispanic origin provide for the collection of data on whether a person is of Hispanic or Latino culture or origin, as defined below.

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 racial groups may be of any of the Hispanic-origin subgroups. The terms “Hispanic or Latino” and “Hispanic” are used interchangeably in this report.

The 1997 OMB guidelines recommend that a two-question format (separate race and Hispanic-ethnicity questions) be used. Most HP2020 data sources that use self-identification for race and ethnicity (for example, National Health Interview and National Health and Nutrition Examination Surveys) used the two-question format. When overriding data collection considerations exist (for example, the Youth Risk Behavior Surveillance System prior to 2007), a combined race-and-ethnicity question that includes a separate Hispanic category equal with the other racial categories can be used.

The 1997 standards on reporting race and ethnicity were implemented by the U.S. Census Bureau starting with the 2000 decennial census. Other federal programs that collect information on race and ethnicity were required to incorporate the new standards into household surveys, administrative forms and records, and other data collections by January 1, 2003. However, not all data sources used in HP2020 adopted these standards over the tracking period, so the format for reporting race and ethnicity differed across some objectives. In the Progress by Population Group analyses, data for all objectives with available race and ethnicity data are provided, although the number of objectives with data for specific racial or ethnic categories varied. Specifically, final progress data were available for the following race or ethnicity categories: American Indian or Alaska Native ($n = 281$ objectives), Asian ($n = 334$ objectives); Native Hawaiian or Other Pacific Islander ($n = 139$ objectives); Hispanic or Latino ($n = 537$ objectives); Black, not Hispanic or Latino ($n = 561$ objectives); and White, not Hispanic or Latino ($n = 575$ objectives). Explanations of how differences in reporting on race and ethnicity across data sources were handled are given below:

- Some HP2020 data sources only reported estimates for Asian or Pacific Islander persons combined rather than as separate Asian and Native Hawaiian or Other Pacific Islander populations, as specified by the 1997 guidelines. In the Progress by Population Group

analyses, objectives that only had data for Asian or Pacific Islander persons combined were included in the Asian category.

- If an HP2020 data source reported estimates for Asian and Asian or Pacific Islander groups separately (but not Native Hawaiian or Other Pacific Islander persons), then only the data for Asian persons were included in measuring final progress (that is, the data for the Asian or Pacific Islander category were not included).
- Data for Mexican-American persons were used for the Hispanic or Latino category if no data were available for the broader Hispanic or Latino category in the data source. If data for both Mexican-American and Hispanic or Latino persons were available for an objective, only the Hispanic or Latino data were included in the Progress by Population Group analyses.
- If race data by Hispanic or Latino ethnicity were not available in the data system, the data for Black and White persons overall were included in the analysis in the Black, not Hispanic or Latino and White, not Hispanic or Latino categories, respectively.

Education

Educational attainment is typically measured by the number of years of education a person has completed or by the highest credential received. The following categories for educational attainment were generally used in HP2020 and the Progress by Population Group analyses:

1. Less than high school ($n = 254$ objectives): Persons with less than 12 years of schooling or no high school diploma.
2. High school graduate ($n = 265$ objectives): Persons with either 12 years of schooling, a high school diploma, or GED.
3. Some college ($n = 264$ objectives): Persons with a high school diploma or GED and 13 years or more of schooling, but without a college degree.
4. Associate's degree ($n = 136$ objectives): A.A., A.S., or other 2-year degree.
5. 4-year college degree ($n = 235$ objectives): B.A., B.S., or other 4-year degree.
6. Advanced degree ($n = 171$ objectives): Master's degree, Ph.D, M.D., or other doctoral or advanced degree.

Generally, data on educational attainment are presented for adults aged 25 and over, consistent with guidance from the U.S. Census Bureau (2). However, the ages used to report educational attainment varied by HP2020 data source. For example, the ages used to report educational attainment are 20-44 in the National Survey of Family Growth and 18 and over in the National Survey on Drug Use and Health. Also, some educational attainment categories varied by data system, objective, or both. For example, a data source may have collected data on the category "more than high school attained" without distinguishing the level of education after high school, or "4-year college or more" without distinguishing whether an advanced degree was obtained. In the Progress by Population Group analyses, if the objective's population data on education were not consistent with the categories listed above, the data were classified in the minimum education category for which the data would apply in the analysis measuring final progress. For example, the category "more than high school attained" was included in the "Some college" category, while "4-year college or more" was included in the "4-year college degree" category. A full listing of these classifications is as follows:

Education level category included in data source	Classification for Progress by Population Group analyses	Number of objectives (percent of total in final category) that were affected
More than high school	Some college	20 (7.6%)
Some college or more	Some college	2 (0.8%)
Some college or associate's degree (including technical school)	Some college	64 (24.2%)
4-year college or more	4-year college degree	7 (3.0%)
College graduate or more	4-year college degree	46 (19.6%)

Income

Most data sources used in HP2020 defined personal income as money earned before taxes excluding noncash benefits. This would include wages, salaries, rental income, interest, dividends, profits, and fees from their own businesses, pensions, and help from relatives. The sum of all personal incomes in a family unit is referred to as family income, although the operational definition of “family” varied across data sources. Each member of a family is classified according to the total income of the family. Unrelated individuals are classified according to their own individual incomes. Family income data were used to compute the poverty status that is used for many HP2020 objectives.

Poverty status

Family income, adjusted for family size and composition and inflation, is the basis for assigning poverty status to a person. Poverty status measures family income relative to family size and composition (for example, single parent, married couple with children) using the poverty thresholds developed by the U.S. Census Bureau and based on definitions originally developed by the Social Security Administration. These thresholds are updated annually and vary by family size and composition. Detailed poverty thresholds by year, size of family, and number of children are available on the U.S. Census Bureau website (3).

In HP2020, an expanded set of poverty threshold categories beyond that of “above” or “below” were used to evaluate a wider range of inequalities in income distribution and its association with health. The set of family income categories typically used in HP2020 and that was included in the Progress by Population Group analyses is as follows:

1. Poor ($n = 323$ objectives): Below the poverty threshold
2. Near poor ($n = 325$ objectives): At 100%–199% of the poverty threshold
3. Middle income ($n = 324$ objectives): At 200%–399% of the poverty threshold
4. Near-high income ($n = 271$ objectives): At 400%–599% of the poverty threshold
5. High income ($n = 256$ objectives): At or above 600% of the poverty threshold

These categories and thresholds varied somewhat by data source. For example, some data systems only collected family income data relative to specific income ranges, while others used poverty thresholds for income categories that differ from those defined previously. For measuring final progress across HP2020 objectives, some family income ranges were combined.

For example, objectives with a data source that collected data on the income range 400%–499% of the poverty threshold were included in the category “Near-high income: At 400%–599% of the poverty threshold” because this range is a subset of the overall category. Data on income with ranges in a data system that overlapped multiple poverty threshold categories relative to those defined previously were still included in the Progress by Population Group analyses to include as many objectives as possible. The following table shows a listing of these ranges from the original data sources that were not consistent with the categories defined previously, along with information on how these data were assigned in the analyses. If a specific income range was reported, the data were categorized into the above-listed income categories as follows: Less than \$20,000 (Poor), \$20,000-\$34,999 (Near poor), \$35,000 to less than \$50,000 (Middle income), \$50,000-\$74,999 (Near-high income), and Greater than \$75,000 (High income).

Family income level included in data source	Classification for Progress by Population Group analyses	Number of objectives (percent of total in final category) that were affected
≤100% of the poverty threshold	Poor	8 (2.5%)
≤130% of the poverty threshold	Near poor	2 (0.6%)
101%-200% of the poverty threshold	Near poor	8 (2.5%)
Borderline poverty threshold	Near poor	2 (0.6%)
>130% - <350% of the poverty threshold	Middle income	2 (0.6%)
≥200% of the poverty threshold	Middle income	11 (3.4%)
Above threshold	Middle income	2 (0.6%)
≥300% of the poverty threshold	Near-high income	5 (1.8%)
>350% of the poverty threshold	Near-high income	2 (0.7%)
≥400% of the poverty threshold	Near-high income	15 (5.5%)
>500% of the poverty threshold	High income	88 (34.4%)

Note: ≤ indicates less than or equal to, < indicates less than, ≥ indicates greater than or equal to, and > indicates greater than.

Disability

The American Community Survey (ACS) disability questions were adopted as a standard definition of disability for use in HP2020 (4). The six ACS questions used to assess disability are:

For all persons:

- *Is this person deaf or does he/she have serious difficulty hearing?* [Yes/No]
- *Is this person blind or does he/she have serious difficulty seeing even when wearing glasses?* [Yes/No]

For persons aged 5 years and over:

- *Because of a physical, mental, or emotional condition, does this person have serious difficulty concentrating, remembering or making decisions? [Yes/No]*
- *Does this person have serious difficulty walking or climbing stairs? [Yes/No]*
- *Does this person have difficulty dressing or bathing? [Yes/No]*

For persons aged 15 years and over:

- *Because of a physical, mental, or emotional condition, does this person have difficulty doing errands alone, such as visiting a doctor's office or shopping? [Yes/No]*

Some data sources in HP2020 used other definitions of disability, including difficulties with basic actions (a composite measure of disability designed to capture limitations or difficulties in movement, emotional, sensory, or cognitive functioning associated with a health problem) and complex activity limitation (a composite measure of disability that measures inability to function successfully in certain social roles). For data sources that used these definitions, the categories used for disability included “people with either basic or complex” and “people with neither basic nor complex” limitations. These varying definitions of disability were combined to create two categories for use in the Progress by Population Group analyses:

- Persons with disabilities or activity limitations ($n = 154$ objectives)
- Persons without disabilities or activity limitations ($n = 161$ objectives)

Geographic Location

Geographic location in HP2020 may refer to the location of a person's residence, a family's residence, a health care provider, or a school, depending on the focus of the specific objective. Geographic location was typically classified in HP2020 as either metropolitan or nonmetropolitan, and for many data sources these categories were based on standards for metropolitan statistical areas (MSAs) as defined and revised by OMB before each decennial census. An MSA is a county or group of contiguous counties that contains at least one urbanized area of 50,000 or more population. MSAs may also 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 (5).

Alternatively, HP2020 data sources may have instead collected geographic location data categorized as urban or rural areas. Urban areas are within or outside an MSA or within or outside 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, as designated by the U.S. Census Bureau, 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 (6).

Urban–rural data for objectives using the National Vital Statistics System, the National Hospital Ambulatory Medical Care Survey, or the National Ambulatory Medical Care Survey were recalculated in 2017 to reflect the 2013 urban–rural classification system (7).

In the Progress by Population Group analyses, these definitions of geographic location were combined to create two categories:

- Urban or metropolitan ($n = 328$ objectives)
- Rural or nonmetropolitan ($n = 313$ objectives)

Some HP2020 data sources defined geographic location as being in a city instead of an urban–rural location. In the Progress by Population Group analyses, city was classified in the “Urban or metropolitan” category.

Objectives that Met or Exceeded the Target at the Baseline Data Point for Population Groups

A single target value for each HP2020 objective was set based on the total population for the objective. Targets were not set for separate population groups. Consequently, the magnitude of the change in a population group estimate during the HP2020 tracking period that was required to meet or exceed the target value differed by each group. There were also many objectives for which a population group met or exceeded the target already at baseline.

An example of this is provided in the table below for objective *C-1: Reduce the overall cancer death rate*. The target set for the total population, 161.4 deaths per 100,000 population (age adjusted to the year 2000 standard population), had already been exceeded for three race or ethnicity groups (American Indian or Alaska Native persons, Hispanic or Latino persons, and Asian or Pacific Islander persons) at the baseline (2007). The baseline year rate for the other two race or ethnicity groups (Black, not Hispanic or Latino persons and White, not Hispanic or Latino persons) had not yet met or exceeded the target set for the total population, but the movement towards the target required to meet this threshold was higher among Black, not Hispanic or Latino persons compared with White, not Hispanic or Latino persons. By the final year of the HP2020 tracking period for this objective (2017), the rate for White, not Hispanic or Latino persons ultimately exceeded the target, whereas the rate for Black, not Hispanic or Latino persons did not. However, the Black, not Hispanic or Latino population had a larger percentage reduction in the age-adjusted cancer death rate from baseline than the White, not Hispanic or Latino population. This example illustrates why users should interpret the final progress categories for population groups with caution, particularly for objectives in the target met or exceeded progress status category. Note that the HP2020 baseline and final year data for individual population groups and target values are available for all objectives included in the Progress by Population Group analyses as a downloadable link under the chart and table on the HP2020 Final Review website.

Race or ethnicity group	HP2020 target, age-adjusted rate per 100,000 population	Baseline year (2007), age-adjusted rate per 100,000 population	Final year (2017), age-adjusted rate per 100,000 population	Final progress status category
American Indian or Alaska Native persons	161.4	124.1	99.2	Target met or exceeded
Hispanic or Latino persons	161.4	121.8	108.1	Target met or exceeded
Black, not Hispanic or Latino persons	161.4	219.8	178.0	Improved
White, not Hispanic or Latino persons	161.4	182.6	157.9	Target met or exceeded
Asian or Pacific Islander persons	161.4	109.5	94.4	Target met or exceeded

References

1. National Center for Health Statistics. Healthy People 2020 General Data Issues. 2021. Available from: https://www.cdc.gov/nchs/healthy_people/hp2020/hp2020_data_issues.htm.
2. U.S. Census Bureau. Educational attainment. Available from: <https://www.census.gov/topics/education/educational-attainment.html>.
3. U.S. Census Bureau. Poverty thresholds. Available from: <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html>.
4. U.S. Department of Health and Human Services. HHS implementation guidance on data collection standards for race, ethnicity, sex, primary language, and disability status. 2011. Available from: <https://aspe.hhs.gov/datacncl/standards/ACA/4302>.
5. U.S. Census Bureau. Metropolitan and micropolitan. Available from: <https://www.census.gov/programs-surveys/metro-micro.html>.
6. U.S. Census Bureau. 2010 census urban and rural classification and urban area criteria. Available from: <https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html>.
7. National Center for Health Statistics. NCHS urban–rural classification scheme for Counties. 2014. Available from: https://www.cdc.gov/nchs/data_Access/urban_rural.htm.