

NOTE TO READERS

What are the Healthy People 2010 Targets?

Healthy People 2010 is the name of the U.S. government's national health goals and objectives initiative. Beginning with the year 2000, every ten years specific disease prevention and health promotion objectives are developed based on baseline data. The two main goals of this program are: 1) to increase life expectancy and improve quality of health, and, 2) to eliminate health disparities among different population groups. Approximately 467 health objectives are described within 28 categories that cover a wide range of health topics (1).

The Healthy People 2010 objectives were developed by scientists from within and outside the government, and ten health topics were selected as the primary benchmarks of the program (2). These ten Leading Health Indicators include objectives related to physical activity, overweight and obesity, tobacco use, substance abuse, responsible sexual behavior, mental health, injury and violence, environmental quality, immunization, and access to health care.

Several Healthy People 2010 objectives are found throughout this report where a comparable data measure was reported for the city of Boston. For more information on Healthy People 2010, and the development of Healthy People 2020 objectives for the next decade, please visit www.healthypeople.gov.

How do we determine if one percentage (point estimate) is higher or lower than another?

To determine whether two percentages or point estimates are different from one another, one cannot look only at the percentages themselves. One must determine whether the differences between two percentages are “statistically significant.” Statistical significance is a mathematical term used to describe the likelihood that a particular number or rate reflects reality. This term comes into play when researchers measure a particular characteristic of a sample or subset of a group or population, and then apply or infer that result to describe the entire group or population.

In this report one of the data sources cited is a survey called the Boston Behavioral Risk Factor Surveillance Survey (BBRFSS). This is a telephone-based survey in which a number of Boston residents were asked to respond to a series of questions. The entire population of Boston was not called to participate in this survey, as this would not have been feasible or cost-effective. The survey data were collected among a random sample of adults within a random sample of Boston households. The resulting data were applied to represent the entire adult population for the city of Boston and are described as percentage or point estimate. These numeric values are further described by a numeric range called a confidence interval. The confidence interval describes the likelihood that the true percentage estimate falls within the range of values given, and thus describes the error around the percentage estimate. To compare two percentage estimates and determine whether one is higher or lower than the other, one must look not only at the individual percentages but the associated confidence intervals.

For example, in the Chronic Disease section of this report there is a measure of self-reported asthma by Boston neighborhood from the BBRFSS (Figure 14.9). The percentage of Boston residents overall reporting asthma was 10%, while for North Dorchester the percentage was

17%. The confidence interval for Boston was 9.3%-11.6%, and for North Dorchester was 11.4%-23.0%. Although the percentage estimates are numerically different (Boston 10%, North Dorchester 17%), the overlapping confidence intervals indicate that these percentages are statistically similar.

Throughout this report, confidence intervals were calculated for all survey data including data from the Boston Behavioral Risk Factor Surveillance System (BBRFSS), Youth Behavioral Risk Surveillance (YRBS), and Boston Neighborhood Survey (BNS). To determine whether a percentage was higher or lower than another percentage, the confidence intervals were calculated and compared. If the confidence intervals did not overlap, the two percentage estimates were reported as different (one was “higher” or “lower” than the another). If the confidence intervals overlapped, the percentage estimates were reported as similar to one another and no further comparison was made.

What do the terms “insufficient sample size” and “ $n < 5$ ” mean?

In this report the phrase *insufficient sample size* is sometimes used in the Notes, Data Analysis, and Data Resources section to describe data points that are not presented. This particularly occurs when survey data are stratified by population groups and as a result, there is not a large enough sample (number of survey respondents or recorded health events) to allow the presentation of reliable rates. Data are also not presented if a sample size is too low to protect the confidentiality of the respondents.

In addition a notation $n < 5$ is used when there are fewer than five occurrences (for example, births, deaths, new cases of a disease) and thus a rate cannot be presented. There are some instances where combining several years of data can overcome the issue of a sample size that is too small, and therefore allow reporting of those data.

Why do we sometimes combine several years of data?

In certain instances, when there were fewer than five cases or an insufficient sample size in a given year, we combined data from two or more years in order to permit the calculation and presentation of a rate or point estimate. In this report, the title of a chart indicates whether two or more years of data have been combined.

How do we define neighborhood boundaries in this report?

Throughout this report, certain data are presented for individual neighborhoods on a map or bar charts. Prior to presenting data for individual neighborhoods, one must first decide how to define neighborhood boundaries. Zip codes or census tracts can be used to define the geographical boundaries of a neighborhood. These are not the only methods to define neighborhood boundaries. However, zip code and census tract information is often collected along with Boston health outcome data. Therefore, when presenting these data in this report, neighborhood boundaries are defined by either zip codes or census tracts. The majority of bar charts and maps in the report use zip code boundaries to define Boston neighborhoods. The only exceptions are charts and maps that present birth and death data. Census tracts are used to define neighborhoods when presenting birth and death data by neighborhood. Information on the boundaries used to identify neighborhoods can be found in the to the Notes, Data Analysis, and Data Sources section.

Why are some of the data older than other data?

For every section in this report, we tried our best to provide the most recent data available. You will note that some data such as the cancer incidence data from 2005 are older while others are more recent. The data come from various sources. The type of data, the frequency of data collection, the post-collection cleaning and verification process, and resources available to manage and analyze the data play a role determining when data are available.

References

1. **Healthy People 2010 Fact Sheet.** Available at <http://www.healthypeople.gov/About/hpfact.htm> and accessed on 2-22-09.
2. **Healthy People 2010 Leading Health Indicators.** Available at http://www.healthypeople.gov/Document/HTML/uih/uih_4.htm and accessed on 2-24-09.

