

# New York National Comparison Adult Tobacco Survey Overview

Center for Community Health  
Division of Chronic Disease Prevention  
Bureau of Chronic Disease Epidemiology and Surveillance  
Tobacco Surveillance, Evaluation, and Research

## **General Description**

The New York National Comparison Adult Tobacco Survey (NY NATS) was developed by the New York Tobacco Control Program (NY TCP) in partnership with Research Triangle Institute. Similar to the NY Adult Tobacco Survey, the NY NATS is administered to a nationwide sample of the non-institutionalized adult population, aged 18 years or older. Since the data are used as a comparison sample, the NY NATS represents the nation minus New York State. The survey is designed to provide timely surveillance and evaluation data to inform the NY TCP. The NATS assesses (a) adult tobacco-related behavior, attitudes, and beliefs; (b) tobacco purchasing and cessation attempt behavior; (c) health status and health-related problems; (d) attitudes toward and exposure to secondhand smoke; (e) perceptions of risk related to tobacco use; (f) recollection of exposure to tobacco advertising and anti-tobacco health communications; and (g) attitudes toward smoking policies. Some key survey items have been included since inception (e.g., current smoking status); some questions may be included for shorter periods of time to assess time-sensitive issues (e.g., support for tax increases), and other questions may be rotated in and out of the survey instrument as necessary.

## **Data collection methodology**

The NY NATS is a list-assisted random-digit-dial survey. Lists of prescreened numbers that increase the probability of reaching a household are purchased from a private firm. The eligible population is U.S. resident civilians at least 18 years old, in all states except New York State. In the NY NATS, listed households are oversampled. Also, the number of smokers in the sample is increased by oversampling geographical designations (block groups and counties) that have lower percentages of baccalaureate graduates. Sampling weights are created so the sample can be used to make population-level estimates. First, a subject's initial weight is defined as the inverse of the probability of selection. After the initial design weights are calculated, nonresponse adjustments are performed using current census projections based on the 2000 census data. Then adjustments are made for household size and number of telephone lines. Finally, the weights are poststratified to population totals based on combinations of age, race/ethnicity, gender, and geographical region. Respondents were compensated \$20 for completing the survey.

## **Statistical and Analytic Issues**

Unweighted data on the NY NATS represent the actual responses of each respondent, before any

adjustment is made for variation in respondents' probability of selection, disproportionate selection of population subgroups relative to the state's population distribution, or nonresponse. Weighted NY NATS data represent results that have been adjusted to compensate for these issues. Use of the final weight in analysis is necessary if generalizations are to be made from the sample to the population. The procedures for estimating variances described in most statistical texts and used in most statistical software packages are based on the assumption of simple random sampling. However, the data collected in the NY NATS are obtained through a complex sample design; therefore, the direct application of standard statistical analysis methods for variance estimation and hypothesis testing may yield misleading results. There are computer programs available that take such complex sample designs into account. SAS survey procedures, (e.g., SURVEYMEANS and SURVEYREG), SUDAAN, and Epi Info's C-Sample are among those suitable for analyzing NY NATS data, though these are not the only applications available. When using these software products, users must know the stratum (geostr) and the record weight (annualwt): these are on the public use data file. (Some surveys also require a primary sampling unit or PSU; that is not necessary for the NY NATS.) For more information on calculating variance estimations using SAS, see the *SAS/STAT Users Guide*. For information about SUDAAN, see the SUDAAN User's Manual, Release 7.5. For information about Epi Info, see *Epi Info, Version 6.0*.

### **Limitations of data use**

Although the overall number of respondents in the NY NATS is more than sufficiently large for statistical inference purposes, subgroup analyses can lead to estimates that are unreliable. Consequently, users need to pay particular attention to the subgroup sample when analyzing subgroup data, especially within a single data year or geographic area. Small sample sizes may produce unstable estimates. Reliability of an estimate depends on the actual unweighted number of respondents in a category, not on the weighted number. Interpreting and reporting weighted numbers that are based on a small, unweighted number of respondents can mislead the reader into believing that a given finding is much more precise than it actually is. We follow, and recommend that others follow, CDC guidelines of not reporting or interpreting percentages based upon a denominator of fewer than 50 respondents (unweighted sample).