

# Survey Statistics



RTI International is a recognized leader in the development and implementation of innovative procedures for survey sampling, weighting, imputation, and analysis. We have extensive experience designing and analyzing both cross-sectional and longitudinal sample surveys covering a wide variety of populations and substantive issues from drug abuse to education.

## Sampling Design

We specialize in sampling design techniques, including stratification, clustering, multistage and multiphase sampling, composite measures of size, controlled selection, and multiple-frame sampling. Our experience includes constructing and sampling from all types of sampling frames:

- List frames such as businesses and Medicare or Medicaid recipients
- Census-based geographic area frames
- Random-digit-dialing telephone number frames.

## Sampling Rare Populations

Our staff use sampling with probabilities proportional to composite measures of size and screening interviews to efficiently locate and interview members of rare populations. We combine general population screening with list sampling when reasonably complete lists of members of the rare population are available. When necessary, we use network sampling and other innovative sampling methods.

## Design Optimization

We have developed software for determining the optimal sample size and sample allocation that achieves multiple precision constraints at minimum cost. We help our clients formulate their objectives into statistical precision constraints, develop survey cost and variance models, and then determine the sample size and allocation that achieves precision requirements within available resources. It is often necessary to make tradeoffs among different constraints or to adjust constraints to achieve a reasonable level of cost.

## Experimental Design

We embed experimental designs into sample surveys to estimate differences in response rates and measurement errors with alternative survey instruments and/or alternative modes of interviewing. For example, we included a sequential factorial design in the pretest for a major occupational survey. Such pretest results are used to model the effects of experimental factors and their costs to optimize the survey methodology.

## Small Area Estimation

In small area estimation (SAE), statistical models are used to predict estimates for small domains by linking survey outcome variables to a set of known predictor variables. Our expertise in SAE can provide policymakers and researchers with more precise and less costly estimates for disease prevalence and other outcomes for small domains.

## Nonsampling Errors and Nonresponse Bias Analysis

We offer special expertise in nonsampling error, including the evaluation of errors of observation (i.e., measurement and data processing errors) and the evaluation and adjustment for errors of nonobservation (i.e., frame undercoverage and nonresponse). We use a number of analytical techniques to evaluate measurement reliability, validity, and bias, including latent class analysis, structural equation modeling, item response theory, reinterview analysis, and comparison to “gold standards.”

## Weighting

To reduce nonresponse and coverage errors, we weight survey data using weighting class, poststratification, calibration, and raking adjustments. RTI statisticians have done considerable

work in developing and applying nonresponse adjustments using logistic modeling of response propensity. Taking advantage of an iterative proportional fitting algorithm and the logit method, RTI's GEM weighting software reduces nonresponse and coverage bias while minimizing variance inflation due to weighting, including automated trimming and smoothing of extreme weights.

## Imputation

RTI's expertise includes classical hot deck and regression imputation. We have applied our expertise to develop new software and procedures such as

- weighted sequential hot deck imputation that controls the expectation over multiple imputations
- imputations based on a weighted polytomous logistic model for the propensity to respond to a category
- a multivariate imputation procedure using a log-linear model, which can preserve complex associations between responses when some of them are missing
- predictive mean neighborhoods, which combines the best features of hot-deck and model-assisted imputations and can be applied to both discrete and continuous variables.

## Statistical Disclosure Control

RTI has developed an innovative statistical disclosure control methodology called MASSC for creating public use micro data files. MASSC uses optimal survey sampling methods to minimize disclosure risk via subsampling and substitution while controlling information loss.

We also use other techniques to treat data to minimize disclosure risk while maintaining the analytic utility of the data, including deleting identifying variables from a public use file (PUF), data coarsening, and data swapping.

---

### For more information, please contact

Roy W. Whitmore, PhD  
Vice President, Statistics and Epidemiology  
Phone: 919.541.5809      E-mail: rww@rti.org

RTI International  
PO Box 12194, 3040 Cornwallis Road  
Research Triangle Park, NC 27709-2194 USA

---

RTI 5287-2 0109

RTI International is one of the world's leading research institutes, dedicated to improving the human condition by turning knowledge into practice. Our staff of more than 2,800 provides research and technical expertise to governments and businesses in more than 40 countries in the areas of health and pharmaceuticals, education and training, surveys and statistics, advanced technology, international development, economic and social policy, energy and the environment, and laboratory and chemistry services. For more information, visit [www.rti.org](http://www.rti.org).

RTI International is a trade name of Research Triangle Institute.

