

II. Designing A Community HANES

This section explores critical decisions to be made early in the design and planning stages of a community HANES. The choices made at this stage serve as the blueprint for the entire survey.

A. Determine Key Goals, Objectives and Subject Area Priorities

1. What does the agency already know, and what are the limitations of these data?

Many states and cities compile statistics on leading causes of morbidity and mortality, such as diabetes, obesity, heart disease, asthma and/or infectious diseases. Health behaviors such as sexual practices or cigarette smoking are also commonly tracked. Thinking about the types of morbidity and mortality data already available within your organization will help determine your study objectives.

Like most local health agencies, NYC uses death certificates and hospital records to gauge trends in morbidity and mortality. Other information on self-reported conditions are also collected through an annual population-based, telephone survey. While these sources provide useful information on various health conditions, there are some important limitations. For example, undiagnosed conditions are not reported and often there is poor recall of health conditions and behaviors. In addition, self-reported blood pressure, cholesterol level and other health measures may be inaccurate, and truthful responses to sensitive questions can be hard to obtain.

2. What does the agency want to find out?

A review of the available statistical data and survey results can help to identify specific health conditions in the community that merit investigation. For example, planners had some insight into the burden of cardiovascular disease and diabetes from hospitalization records and death certificates. However, what was not known was how many people suffered from or were aware that they had these conditions, or were currently undergoing treatment. Having information on the prevalence, awareness

and treatment of chronic diseases and their precursors, as determined through objective measurements, can now help the chronic disease programs target their outreach and public information efforts.

3. What conditions can only be evaluated via a physical exam, and/or a biological specimen?

Nationally, NHANES provides unique and valuable information on health conditions that require:

- Physical examination (e.g., blood pressure, height and weight)
- Biologic specimen testing (e.g., fasting plasma glucose and blood lipid profile)
- Information on sensitive conditions that are not easily ascertained by a brief telephone survey (e.g., mental health and sexual behavior)

Because it is more cost-effective to collect non-sensitive, self-reported data via other approaches (such as telephone interviews), HANES organizers may wish to concentrate on conditions that are best measured by a clinical examination, laboratory analysis and private and confidential questionnaire methods.

For example, if an agency wants to explore the level of mercury exposure among adults in the community, surveillance data or telephone interviews would not be very useful, since this information is collected and reported only for a select number of people who are sufficiently concerned or knowledgeable about mercury. And only those who have been recently tested will likely recall their test results. The most accurate way to assess community mercury exposure would be to conduct biological testing on a representative sample.

[View: NYC HANES data & research questions

http://www.nyc.gov/html/doh/downloads/pdf/hanes/HANES_design_dandr_questions.pdf]

B. Solicit Input

As organizers begin to think about what conditions to measure, it is important to solicit input from programs within the agency that will be using the data. The first step is to identify programs within the agency that should be involved. Their support and cooperation will be essential to the study's success:

Two basic questions to ask during planning are:

- What do these agency programs want to know?
- What do they believe are the best methods for collecting this information?

C. Determine the Feasibility

After creating an inventory of study topics and measures, study organizers should determine how best to obtain each piece of information. These methods will include ascertaining information through survey questions, conducting physical assessments and/or collecting blood

and urine for biological testing. The study planners' task is to determine which of these methods may be best suited to answer each question, and whether it can realistically be accomplished. For example, asthma can be assessed via lung function testing. However, such tests require specialized equipment and skilled staff. Alternatively, obesity is evaluated via a simple physical examination that measures height and weight, and central adiposity. Striking a balance between the desire to have a physical measurement and the practical and logistical costs of obtaining it is an important consideration.

In determining the best way to measure conditions of interest, study planners should always strive for standardization with NHANES' methods. Comparability between national rates and community-level rates can only be achieved if conditions are measured using standard protocols. NYC HANES planners worked closely with NCHS staff during the study design phase to ensure comparability.

DOHMH Programs Involved in Conceptualizing NYC HANES

Office of the Commissioner

Division of Epidemiology
Bureau of Epidemiology Services

Division of Health Promotion and Disease Prevention
Bureau of Chronic Disease Prevention
Bureau of Tobacco Control and Prevention

Division of Environmental Health Services
Environmental Public Health Tracking Program
Environmental & Occupational Disease Epidemiology

Division of Disease Control
Public Health Laboratories
Bureau of Sexually Transmitted Disease Control

Division of Financial & Strategic Management
Bureau of Communications

Pesticide Exposure in New York City

DOHMH is a part of a national effort led by CDC's National Center for Environmental Health (NCEH) to build an Environmental Public Health Tracking Network (<http://www.cdc.gov/nceh/tracking/>). NYC's work involves creating a pesticide tracking system to evaluate the extent of pest infestation and risks to health from pests as well as pest control.* Though substantial data is available on pesticide use from several sources such as a statewide registry of reports of pesticide applications, telephone surveys conducted by DOHMH, and the agency's poison control center, exposure data was still lacking.

NYC's Environmental Public Health Tracking staff worked alongside NYC HANES planners to incorporate an evaluation of pesticide exposure into the physical examination component of the survey. Working with the NCEH laboratories at CDC, planners determined that for relatively little additional cost, urine samples could be collected and analyzed for metabolites of two important classes of pesticides commonly used for structural pest control in cities – organophosphates and pyrethrins/pyrethroids. Collecting urine samples also enabled NYC HANES to evaluate urinary metals analytes and to create an additional specimen repository for future analysis.

These analyses will permit NYC to answer several key questions:

- How does an urban population's pesticide exposure differ from that of the nation as a whole?
- Do exposure profiles differ among major demographic sub-groups?
- What characteristics are associated with higher and lower exposures?
- Does exposure to banned pesticides persist?

*Kass D, Thier A, Leighton J, Cone J, Jeffery N (2004). "Developing a comprehensive pesticide health effects tracking system for an urban setting: New York City's experience. *Environmental Health Perspectives*. 112(14): 1419-23.

HIV Testing and NYC HANES

In the initial planning phases, NYC HANES planners considered conducting testing for HIV as part of the main study. However New York state law requires that persons tested for HIV undergo pretest counseling and that those who test positive for HIV undergo a face-to-face post-test counseling interview. This type of lengthy counseling represented a logistical challenge for planners, as they wanted to minimize both the time burden on the participant and the invasiveness of the study components, so as not to impact response rates. Ultimately, the decision was made to not include HIV testing in the NYC HANES protocol.

Health Measures Assessed in NYC HANES

NYC HANES planners worked closely with NCHS staff to ensure comparability with NHANES data collection protocols for all aspect of the survey.

Medical exams

- Blood pressure (systolic and diastolic)
- Height, weight, arm and waist circumference
- Venipuncture and urine collection

Blood tests

- Lipid profile
- Diabetes measures (fasting plasma glucose, glycohemoglobin)
- Cotinine (a by-product of tobacco smoke)
- Hepatitis C virus
- Herpes simplex 2 virus
- Mercury, cadmium and lead

Urine tests

- Exposure to pesticides and metals

In-person interviews

- Health status questions
- Nutrition questions
- Tobacco use
- Reproductive health
- Mental health

Audio Computer Assisted Self Interview (ACASI)

- Sexual experience
- Drug and alcohol use

D. Identify the Target Population

Determining your target population is the first and most important step in designing any survey. The target population may be evident from the priorities within your agency. The final study component list and feasibility factors will also influence your choice. For example, if your research goals involve learning more about the diseases and health conditions that impact children, your target population may be limited to persons under 18 years old.

Time and budget considerations will also directly impact the number of participants that can be

recruited into any study. Planners will need to narrow their focus to identify the target population that will best match the research goals.

NYC HANES was designed as an adult survey to assess adult risk behaviors and conditions; therefore the target population consisted of adults aged 20 and older. Each community will need to determine its' population of interest. Planners wishing to include children will need to consider additional factors such as parental consent and logistical challenges involved in interviewing and examining this population.

E. Determine Optimal and Feasible Budgets

Whether conducted at the state or community level, a HANES can be a large, expensive project. In the planning stages several different strategies may need to be considered to arrive at the most cost-effective study design.

1. What are the major costs to consider?

The first step in planning a HANES budget is to determine the major cost categories that are required to implement the survey plan. For NYC HANES, major costs included:

- Staff
- Facilities and equipment
- Sample design and selection
- Data collection systems
- Materials
- Transportation
- Participant remuneration
- Laboratory analysis

2. How much will it cost to collect the data?

The next step is to estimate how much money will be needed to fund each major category. For example, if new staff members are to be hired, the number and type of staff will need to be estimated, as well as the associated salary costs. If your study design requires a large sample, this will affect the costs associated with materials and laboratory analyses. Also consider whether remuneration will be provided as an incentive for recruitment, and if transportation to clinic facilities will be provided.

3. How much money is available and from where?

Study organizers should have an initial idea of how much money is available from departmental funds or special project allocations. If local funding is not available or is not sufficient, there may be state, federal or private grant funding available. Planners may also want to consider partnering with academic institutions that can offer critical intellectual, staff and monetary resources.

NYC HANES was financed primarily through city funding, however several components were funded from other sources. For example, grant funds were used to finance the environmental bio-monitoring component (metals and pesticides) while the New York State Department of Health supported the environmental tobacco smoke component.

4. How can the survey design and operations be tailored to fit a specific budget?

As often happens, the funds available to implement a project may be lower than the projected costs. To reduce costs, survey organizers should consider maximizing existing resources and relationships. To arrive at the most cost-effective plan, NYC HANES planners tried to take advantage of currently available equipment, supplies, physical facilities and DOHMH personnel. Although some equipment and space were borrowed from other parts of the agency, extensive purchasing of specialized equipment and supplies was still required.

NYC HANES' Target Population

New York City's population is a heterogeneous mix of ethnic and socioeconomic groups throughout five boroughs (The Bronx, Brooklyn, Manhattan, Staten Island and Queens). To adequately sample this broad mix, the NYC HANES targeted adults in every borough and included non-English speakers, illiterate individuals, pregnant women, and the mentally or developmentally disabled. However, adults living in group quarters such as college dormitories, homeless shelters, or nursing homes were excluded.

F. Develop the Sampling Design

After decisions regarding the target population and budget have been made, the next step involves developing the sampling design. Important considerations include determining the sample size that will be needed and deciding how participants will be selected into the study.

1. Estimate the Required Sample Size

The size of your target sample will depend on the total population, the expected disease and exposure prevalence in that population, your estimated response rate, and the budget available. The goal is to choose a sample size that provides sufficient statistical power to accurately estimate the prevalence of diseases of interest that can be evaluated given your budget, staff and time limitations.

The NHANES sampling frame is developed through contract with Westat, a private consulting firm. In consultation with NYC HANES project leaders Westat constructed the NYC HANES sample design. Using NHANES sample size calculations, as well as locally available data, Westat calculated the sample needed to produce reliable NYC estimates for conditions of interest, such as diabetes and hypertension. A minimum target sample size for NYC HANES of 2,000 participants was set in order to ensure enough statistical power to estimate important health conditions in NYC. Because of the diversity of the New York City population, there was no need to over-sample different demographic groups.

2. Estimate the response rate

Once the sample size has been determined, you will need to estimate how many households will need to be contacted in order to recruit enough participants to achieve your required sample size. Response rates reflect the degree of success you will have in contacting selected households, completing a short eligibility interview, and enrolling selected survey participants into the study.

To get an idea of the expected response rate, it may be possible to assess response rates from similar surveys conducted in your community. It will be important to draw from the experience of programs that have conducted population surveys, or university researchers that may have carried out similar studies. If the expected local response rate is low, this can introduce bias into the study findings and potentially threaten the study's statistical power and accuracy. In this case organizers may want to consider increasing the target sample size. Several factors are likely to affect a response rate:

- **Staff training and skills:** Well-trained staff members are better equipped at reaching residents of selected households, using screening instruments properly, and persuading selected residents to participate.
- **Vacancy rate:** A certain proportion of households will be vacant, influencing the number of households that need to be approached. This information is available from the U.S. Census Bureau. Vacancy rates can also be estimated from American Housing Survey and American Community Survey data. In NYC a triennial Housing and Vacancy Survey is conducted that provides neighborhood-scale vacancy rates.
- **Appointment availability:** The availability of early morning, evening and weekend appointments for both the clinic setting and/or home visits can increase response rates.
- **Burden of participation:** The amount of time required to complete the interview and exam, and how far a participant must travel will also directly affect response rates.
- **Incentives:** Providing remuneration, a prize or other incentives, such as transportation, to participants can increase the response rate.

NYC HANES Response Rate and Estimates

For several reasons NYC HANES planners anticipated lower response rates than those attained in NHANES. Experience had shown that response rates in Northeastern urban cities tend to be lower than in other parts of the country. Additionally, NYC HANES field staff would be less experienced overall in survey techniques than NHANES staff.

The study design team assumed that:

- 8 percent of all households approached would be vacant.
- 80 percent of selected households would successfully complete an initial screening questionnaire used to determine eligibility.
- 75 percent of selected eligible survey participants would successfully complete the interview and examination, yielding a
- 60 percent overall response rate.

Based on these estimates, NYC HANES was designed to target:

- 4,082 dwelling units, to yield
- 3,004 screened households, to obtain
- 2,657 subjects and
- **2,000 participants** completing the full examination.

3. Develop a Sampling Plan

The goal of the NYC HANES sampling plan was to obtain a representative sample such that every eligible person in the population of interest had an equal chance of being selected. The best way to achieve a representative sample is to randomly select members of the population. In most situations this may not be a feasible option due to the cost and logistical difficulty of preparing a representative list of all residents. The commonly used alternative approach is a multi-stage probability sample. NHANES uses a four-stage probability sample to obtain a representative sample of the U.S. population. For NYC HANES, a three-stage cluster-sampling plan was used.

G. Develop a Data Collection Model

Once the sampling plans are in place, the survey needs to be operationalized. While operational plans will vary depending on geographic area and the characteristics of the community, some important issues should be considered:

- a) How will the staff approach households to determine eligibility? Will they go door-to-door, obtain telephone numbers and call, or contact residents by mail? (A combination of methods could be used.)
- b) Where will identified survey participants go to complete the survey? Will interviews and exams be done in their home, at a fixed-site facility or a mobile facility?

NYC HANES Sample Design

Stage 1: Select geographic segments

- The geographic segments, consisting of a block or group of proximal blocks within a given NYC census tract, were selected based on probability proportional to size (PPS).
- 144 geographic area segments were randomly selected across the five boroughs of the city.

Stage 2: Select households

- All households in a selected geographic segment were enumerated.
- A random sample of households was generated from all enumerated households in the segments. This included the targeted number of households and a reserve sample to be used if response rates were below target and the sample size was too small.
- A sample of 5,699 households (including the reserve sample) was randomly selected from the 144 segments.

Stage 3: Select individuals

- Adults within households were selected for inclusion in the study.
- Eligible adults age 20 and older were randomly selected based on an *a priori* computer-generated sampling flag.
- The adult sampling procedure in NYC was designed to select either zero, one, or two adults from each selected household, depending on the total number of adults residing in that unit.
- To be sure that all adults had the same probability of being selected, sampling rates were varied based on the number of people in each household.
- 2000 adults were targeted for selection (the final number depended on the household composition encountered).

[View: NYC HANES sample design description

http://www.nyc.gov/html/doh/downloads/pdf/hanes/HANES_design_sample_description.pdf]

c) Given the data that is being collected, what equipment and space will be required?

d) What staffing model is right for the survey? Should existing staff be used, new staff hired, or will outside staff be contracted, either individually or through a business entity?

e) How will data be captured? On paper, electronically or a combination of the two? What are the existing data capacities within the agency?

f) Given answers to the above questions, what are the discreet groups of people/jobs required?

NYC HANES Operational Model

NYC HANES operations took place in three distinct operational groups:

1. Field Operations – Recruitment staff approached and screened selected households to determine participant eligibility.
2. Clinic Operations – Clinic staff working at fixed-site facilities conducted interviews and exams.
3. Central Office Operations – The central office served as the coordinating center for all NYC HANES operations. The staff’s main functions included scheduling and facilitating appointments; processing household visits and follow-up information; handling data management; and coordinating communication among staff.

H. Establish the Timeline

Once key decisions about the data collection model have been made, the next step is to develop an overall timeline for the project. Discrete phases that will require allocated time may include:

- Determining the sample design
- Developing the survey content
- Preparing the study protocol & Internal Review Board (IRB) review
- Staff recruitment and training
- A staff dress rehearsal
- Data collection
- Data preparation
- Data analysis and dissemination

As you establish the schedule for your community HANES, remember to include some leeway for delays and unanticipated problems.

NYC HANES Timeline

ID	Task Name	Start	Finish	Duration	Q1 03		Q2 03		Q3 03			Q4 03			Q1 04			Q2 04			Q3 04			Q4 04			Q1 05			Q2 05
					Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
1	PHASE I: STUDY DEVELOPMENT	2/3/2003	12/30/2003	47w 2d	[Green bar spanning from Feb 2003 to Dec 2003]																									
2	Solidify proposal with consultant	2/3/2003	7/15/2003	23w 2d	[Green bar from Feb 2003 to Jul 2003]																									
3	Determine sampling design	5/1/2003	7/30/2003	13w	[Green bar from May 2003 to Jul 2003]																									
4	Prepare study protocol	7/15/2003	8/29/2003	6w 4d	[Green bar from Jul 2003 to Aug 2003]																									
5	IRB protocol review	9/1/2003	4/29/2004	34w 4d	[Green bar from Sep 2003 to Apr 2004]																									
6	Content development	2/3/2003	9/30/2003	34w 2d	[Green bar from Feb 2003 to Sep 2003]																									
7	Application development & implementation	6/2/2003	12/30/2003	30w 2d	[Green bar from Jun 2003 to Dec 2003]																									
8	PHASE II: PILOT TESTING	1/15/2004	4/30/2004	15w 2d	[Green bar from Jan 2004 to Apr 2004]																									
9	Pilot Phase 1: Mock interviewing	1/15/2004	2/13/2004	4w 2d	[Green bar from Jan 2004 to Feb 2004]																									
10	Pilot Phase 2: Revisions	3/1/2004	3/22/2004	3w 1d	[Green bar from Mar 2004 to Mar 2004]																									
11	NYC deployment	3/29/2004	4/12/2004	2w 1d	[Green bar from Mar 2004 to Mar 2004]																									
12	Staff training	5/11/2004	5/20/2004	1w 3d	[Green bar from May 2004 to May 2004]																									
13	Dress rehearsal	5/14/2004	6/2/2004	2w 4d	[Green bar from May 2004 to May 2004]																									
14	PHASE III: DATA COLLECTION	6/16/2004	12/17/2004	26w 3d	[Green bar from Jun 2004 to Dec 2004]																									
15	PHASE IV: DATA PREPARATION	12/20/2004	4/11/2005	16w 1d	[Green bar from Dec 2004 to Apr 2005]																									
16	PHASE V: DATA ANALYSIS & DISSEMINATION	3/1/2005	on going	on going	[Green bar from Mar 2005 to Apr 2005]																									