

## THE COMMUNITY TRACKING STUDY: MEASURING CHANGES IN THE U. S. HEALTH CARE SYSTEM OVER TIME

Barbara Lepidus Carlson, Richard Strouse, and John Hall  
Mathematica Policy Research, P.O. Box 2393, Princeton, NJ 08543-2393

**KEY WORDS:** Community Tracking Study, health care, health survey

**Abstract.** The Center for Studying Health System Change (HSC), with funding by the Robert Wood Johnson Foundation, has undertaken a multi-year study of changes over time in the U.S. health care system. While the Community Tracking Study (CTS) allows for national estimates using survey data, its design also allows for a more focused study of twelve randomly selected metropolitan areas, including data from site visits. There are three survey components in the CTS: a household survey, a physician survey, and a followback survey of health plans. To date, three rounds of data collection have been completed for the household and physician surveys, and two rounds for the followback survey. Data are made available to researchers via public use files. More information about the Community Tracking Study is at <http://www.hschange.com/>.

To allow for a number of different types of survey estimates while maintaining an efficient design, the sample designs are complex. The design characteristics that make the data so versatile have posed many challenges in the estimation process.

**Goals of the Community Tracking Study.** The CTS is designed to provide a sound information base for decision making by health care leaders. It does so by collecting information on how the health system is evolving in 60 nationally representative communities across the United States and on the effects of those changes on people. The CTS, which has been under way since 1996, is a longitudinal project that relies on periodic site visits and surveys of households, physicians, and employers. The CTS addresses two broad questions that are important to public and private health decision makers:

1. How is the health system changing? How are hospitals, health plans, physicians, safety net providers, and other provider groups restructuring, and what key forces are driving organizational change?

2. How do these changes affect people? How are insurance coverage, access to care, use of services, health care costs, and perceived quality of health care changing over time?

Focusing on markets is central to the design of the CTS. Understanding market changes requires a study

of local markets, including the markets' culture, history, and public policies relating to health care. To track change across the United States, we randomly selected 60 nationally representative communities<sup>1</sup> stratified by region, community size, and type (metropolitan or non-metropolitan). The design also allows researchers to make valid national estimates using the data.

The CTS examines 12 of the 60 communities in depth by conducting site visits and using survey samples large enough to draw conclusions about change in each community. The 12 communities, referred to as *high-intensity sites*, comprise a randomly selected subset of sites that are metropolitan areas with more than 200,000 people (as of July 1992). These sites are: Boston (MA), Cleveland (OH), Greenville (SC), Indianapolis (IN), Lansing (MI), Little Rock (AR), Miami (FL), Newark (NJ), Orange County (CA), Phoenix (AZ), Seattle (WA), and Syracuse (NY).

**Analytic Components of the CTS.** The CTS has qualitative and quantitative components. The qualitative component consists of bi-annual case studies with health policy decision makers in the 12 high-intensity sites. Complementing this information are the CTS surveys.

The CTS includes independent surveys of households, physicians, and employers in all 60 sites, thereby enabling researchers to explore relationships among purchasers, providers, and consumers of health care at the site level.<sup>2</sup> Since 1997, three

---

<sup>1</sup> The CTS covers the contiguous 48 states and the District of Columbia. Alaska and Hawaii were not part of the study.

<sup>2</sup> Under HSC's direction, Mathematica Policy Research (MPR) designed the household and physician survey samples and weights. MPR conducted the household and followback surveys and Gallup collected the physician surveys. Final data processing and file production were carried out by Social and Scientific Systems. The RAND Corporation, in collaboration with HSC, conducted the Employer Survey. RTI did the data collection for that survey.

rounds of the household and physician surveys have been completed; the employer survey was only conducted for the first round. The next round of the household survey is planned for 2003, and the next physician survey is scheduled for 2004.

An insurance followback survey, which is linked to the household survey, was conducted for the first two rounds. In the followback survey, the privately financed health insurance policies covering respondents to the survey of households are “followed back” to the organization that administers the policy. The purpose of this survey is to obtain information about the private policies that is more detailed and more accurate than household survey respondents are able to provide.

The 60 CTS sites were selected with probability proportionate to size (based on population in 1992), stratified by region, community size, and metropolitan status. Forty-eight of the sites were selected from among “large” metropolitan statistical areas (MSAs); that is, with a population of 200,000 or more. Three sites were selected from among small MSAs, and nine from non-MSAs. The boundaries for these non-metropolitan markets were loosely based on Bureau of Economic Analysis (BEA) boundaries. An independent national sample was selected to improve precision of national estimates.

**Household Survey.** All household survey interviews are conducted via telephone using computer-assisted telephone interviewing (CATI).<sup>3</sup> The initial part of the interview uses a household informant to enumerate all household members, their ages, and their relationships to the householder. From this information, the CATI program forms family insurance units (FIUs) within each household. An FIU links adults to their spouses and to their minor children (if any). These units are designed to represent people who can generally be covered together under a typical family health insurance policy. Once the FIUs are formed, the interview continues separately for each FIU, with one respondent per FIU answering questions about each

---

<sup>3</sup> The software used for the household and followback CATI instruments was CASES, a UNIX-based system developed by the Computer-Assisted Survey Methods (CMS) Program at the University of California, Berkeley. (Neither the CSM staff nor the University of California bear any responsibility for the results or conclusions presented here or elsewhere.) The SURVENT system was used to program the physician survey CATI instruments.

FIU member. If an FIU has more than one child, the CATI program randomly selects one child to ask questions about. In addition, each adult in the FIU is asked to respond to questions about health status and other subjective measures.

The sample of households for this survey is comprised primarily of telephone numbers selected using list-assisted random digit dialing (RDD). Because this method includes only those who can be reached by telephone, we have supplemented the RDD sample with an area probability sample in the twelve high-intensity sites. These additional households have either no telephone service or had an interruption in telephone service in the past year. The field interviewer conducts a screening interview to determine telephone status, and then facilitates the interview, which is conducted via cellular telephone by an interviewer at one of MPR’s telephone operations centers. After round one, we included in our samples a certain proportion of telephone numbers (and addresses) that had been part of the prior round’s sample—oversampling those who had completed the interview in the prior round. New sample supplemented this “overlap” sample in each subsequent round.

The main topics covered by the household survey are:

- Household composition
- Health insurance coverage
- Use of health services
- Unmet needs and expenses
- Usual source of care
- Patient trust and satisfaction
- Last physician visit
- Health status
- Chronic diseases
- Risk behaviors
- Employment and earnings
- Demographics

In round three, the household survey had data on 59,725 people (1,370 of these from the in-person sample) in 32,669 FIUs (925 of these in-person). The overall unweighted response rate in round three was 61 percent and the weighted response rate was 59 percent).

**Physician Survey.** The physician survey is designed to document changes that allopathic (MD) and osteopathic (DO) physicians are experiencing in the health care system and to learn how these changes are affecting physicians, their practices, and the way they deliver medical care to their patients. The goal is to provide information to public and private leaders that will enable them to make better policy decisions.

Some of the analytic areas include:

- Impact of managed care participation on physician behavior, perceptions of quality of care provided and physician satisfaction.

- Effects of physician practice arrangements, ownership and risk-bearing on the practice of medicine.

- Relationships between the distribution of practice revenue and physician practice style and satisfaction.

- Effects of socio-demographic or market factors on physicians' practice revenues or income.

- Impact of federal, state and local policies affecting physician practice (including Medicare and Medicaid policy) on physician behaviors and perceptions of their impact on quality of care.

The survey is a nationally representative telephone survey of non-federal, patient care physicians. Each round of the Physician Survey contains observations from more than 12,000 physicians who spend at least 20 hours a week in direct patient care. Approximately 90 percent of the interviews were collected from physicians practicing in the 60 CTS sites; the remaining 10 percent were with physicians selected from a nationally representative supplement designed to improve national estimates.

In round three, the physician survey included 12,406 interviews (7,673 primary care physicians and 4,733 specialists). The round three response rates were the same as for the household survey—61 percent unweighted and 59 percent weighted.

**Insurer Followback Survey.** The followback survey is linked to the household survey. We start with household survey respondents, under age 65, who reported having private health insurance coverage. We then attempt to contact the insurance carrier (and employer if necessary) to obtain information about the insurance product mentioned by the household survey respondent.

This survey is limited to households within the boundaries of the 60 sites. Data collection methods differed between rounds one and two. During the first round, information was obtained by a mix of telephone interviews and faxed forms provided by insurers. For round 2, data were obtained by CATI.

The goal of the followback was to match household reported insurance products to the insurance products described by the insurers. Uniquely matched insurance products were called “hard matches.” If we were able to match a product to one offered by an insurance carrier within a CTS cite, but not to a specific product, we called these links “soft matches.” To assign soft matches to a product, we used a probability matching technique suggested by A. C. Singh *et al.*<sup>4</sup> to find the most likely match.

The main topics in the followback survey are:

- Product type
- In-network and out-of-network coverage
- Provider payment methods
- Consumer cost-sharing

For each of the two rounds for which we conducted an insurer followback, we were able to hard-match approximately half of the products—53 percent for round one and 51 percent for round two. Soft match rates were 19 percent for round one and 9 percent for round two.

**Public Use Files.** HSC makes CTS data available to researchers via public use files and restricted use files. The restricted use files provide more detailed data than the public use files; however, they also require the researcher to sign a data use agreement, which protects the confidentiality of respondents. This agreement covers ownership of data, when to access data and by whom, how data may be used, data security procedures to implement, and sanctions for data misuse.

The household survey public use file allows for national or site-level estimates of the data. Unlike the restricted use version of the file, the public use file does not allow for county-level estimates. Furthermore, the public use file contains no data from the followback survey, which is available on the restricted use file; several variables have been deleted or modified, compared to the restricted use file; and linking the data for specific persons or families across rounds is possible only with the restricted use file.

To protect the confidentiality of our physician respondents, the physician survey public use file

---

<sup>4</sup> Singh, A.C., H.J. Mantel, M.D. Kinack, and G. Rowe. “Statistical Matching: Use of Auxiliary Information as an Alternative to the Conditional Independence Assumption.” *Survey Methodology*, vol. 19, 1993, pp. 59-79.

allows only for national estimates, whereas the restricted use file also allows for site-level estimates. There are no market area (site) identifiers on the public use file. HSC does provide, however, a summary file containing market-level means for most of the variables in the physician survey. As with the household survey public use file, several variables have been deleted or modified, compared to the restricted use file. Similarly, linking the data for specific physicians across rounds is possible only with the restricted use version of the file.

Due to the complex sample design employed in the CTS, specialized software (such as SUDAAN) is required to correctly calculate the variance of estimates. Because the variance estimation parameters needed to run this software includes site identifiers, the physician survey public use file does not include such parameters. Instead, we supply standard error look-up tables and formulas so that researchers can estimate the correct variances.

With each release of a public or restricted use file, there is an accompanying codebook and user guide. The files and documentation are available from the University of Michigan's Inter-university Consortium for Political and Social Research (ICPSR), which can be accessed through the HSC web site: [www.hschange.com](http://www.hschange.com). Technical support for using the data is available at [ctshelp@hschange.org](mailto:ctshelp@hschange.org). The HSC web site also contains numerous Issue Briefs, Data Bulletins, and other reports based on the CTS data, as well as methodological documentation.

***Analysis of CTS Data.*** Analysis weights for the household survey are provided at the person and FIU levels. Analysis weights for the physician survey are provided at the physician level, and for the followback survey are provided at the person level.

As mentioned above, there are two independent national samples in the CTS: a clustered sample within the 60 selected sites (the site sample), and an unclustered national sample (the supplemental sample). The analysis weights provided to make site-specific estimates are based on the site sample cases combined with those supplemental sample cases that happened to fall within the boundaries of the 60 sites.

(The 60 sites represent about 50 percent of the U.S. population.) This combination of sample components is referred to as the "augmented site sample," and maximizes the sample size within each site.

For national estimates, weights are provided for four different configurations of the sample. The preferred configuration is the "combined sample," which combines the site sample with the entire supplemental sample; however, because supplemental sample cases falling outside the boundaries of the 60 sites have no geographic indicators, this configuration would not be useful if a researcher were to include site-level characteristics in his/her model. The alternatives would be to use only the site sample or the augmented site sample in making national estimates. The augmented site sample has a somewhat larger sample than the site sample alone.

A weight is also provided to make national estimates based on the supplemental sample on its own. This smaller sample can be used to perform exploratory analysis and then use the site sample to confirm the results. Because the supplemental sample is unclustered, it has smaller design effects and thus reduces the risk of severely understating the variance inherent in not using more complex statistical packages like SUDAAN to develop variance estimates.

Even with design effects due to clustering and unequal weighting effects, the effective sample size (nominal sample size divided by the design effect) is still large enough to produce reliable national estimates from the CTS.

The tables and figure that follow show the 60 sites, the sample components, and the weights provided for each of the CTS surveys.

TABLE 1. SITES SELECTED FOR THE COMMUNITY TRACKING STUDY

High-Intensity Sites	Low-Intensity Sites		
Metropolitan Areas >200,000 Population	Metropolitan Areas >200,000 Population	Metropolitan Areas <200,000 Population	Nonmetropolitan Areas
01-Boston (MA) 02-Cleveland (OH) 03-Greenville (SC) 04-Indianapolis (IN) 05-Lansing (MI) 06-Little Rock (AR) 07-Miami (FL) 08-Newark (NJ) 09-Orange County (CA) 10-Phoenix (AZ) 11-Seattle (WA) 12-Syracuse (NY)	13-Atlanta (GA) 14-Augusta (GA/SC) 15-Baltimore (MD) 16-Bridgeport (CT) 17-Chicago (IL) 18-Columbus (OH) 19-Denver (CO) 20-Detroit (MI) 21-Greensboro (NC) 22-Houston (TX) 23-Huntington (WV/KY/OH) 24-Killeen (TX) 25-Knoxville (TN) 26-Las Vegas (NV/AZ) 27-Los Angeles (CA) 28-Middlesex (NJ) 29-Milwaukee (WI) 30-Minneapolis (MN/WI) 31-Modesto (CA) 32-Nassau (NY) 33-New York City (NY) 34-Philadelphia (PA/NJ) 35-Pittsburgh (PA) 36-Portland (OR/WA) 37-Riverside (CA) 38-Rochester (NY) 39-San Antonio (TX) 40-San Francisco (CA) 41-Santa Rosa (CA) 42-Shreveport (LA) 43-St. Louis (MO/IL) 44-Tampa (FL) 45-Tulsa (OK) 46-Washington (DC/MD/VA) 47-West Palm Beach (FL) 48-Worcester (MA)	49-Dothan (AL) 50-Terre Haute (IN) 51-Wilmington (NC)	52-West Central Alabama 53-Central Arkansas 54-Northern Georgia 55-Northeastern Illinois 56-Northeastern Indiana 57-Eastern Maine 58-Eastern North Carolina 59-Northern Utah 60-Northwestern Washington

FIGURE 1. THE COMMUNITY TRACKING STUDY SAMPLE STRUCTURE

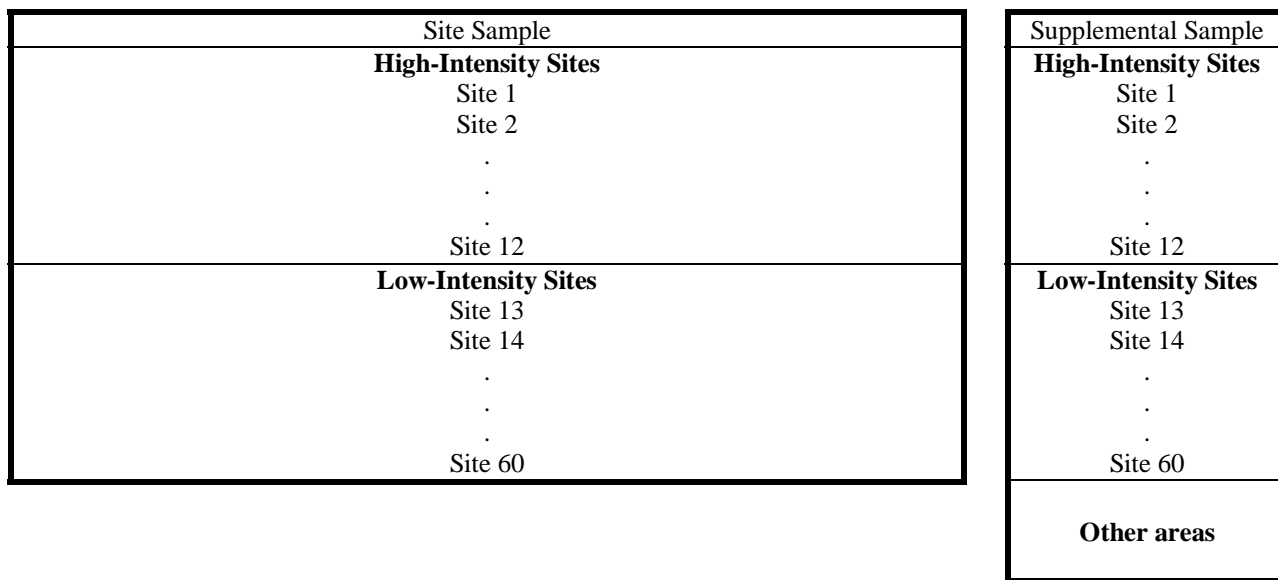


TABLE 2. WEIGHTS PROVIDED ON PUBLIC AND RESTRICTED USE FILES

Type of Estimate	Sample(s)	Household Survey (PUF and RUF)		Followback Survey (RUF)	Physician Survey	
		Person-level	FIU-level	Person-level	RUF	PUF
Site-specific	Augmented site sample	WTPER1	WTFAM1	FBWTPER1	WTPHY1	N/A
<b>National:</b>						
Model includes site characteristics	Site sample	WTPER2	WTFAM2	N/A	N/A	N/A
	Augmented site sample	WTPER5	WTFAM5	N/A	WTPHY5	N/A
	Panel sample	N/A	N/A	N/A	WTPAN2	N/A
Model includes followback variables	Augmented site sample	N/A	N/A	FBWTPER5	N/A	N/A
No site characteristics or followback variables in model	Supplemental sample	WTPER3	WTFAM3	N/A	WTPHY3	N/A
	Combined sample	WTPER4	WTFAM4	N/A	WTPHY4	WTPHY4
	Panel sample	N/A	N/A	N/A	WTPAN1	N/A