Community Tracking Study Household and Physician Surveys

SHADAC Data Information Session
March 10, 2004

Elizabeth Schaefer, *Data Projects Manager*
Center for Studying Health System Change
600 Maryland Ave., Suite 550, Washington, DC 20024
www.hschange.org
eschaefer@hschange.org; 202-484-4232

Overview of the Community Tracking Study (CTS)

The Community Tracking Study (CTS) is conducted by the Center for Studying Health System Change (HSC) and funded by the Robert Wood Johnson Foundation.

- HSC is a nonpartisan policy research organization located in Washington, D.C.
- The purpose of the CTS is to study how the health system is changing (both locally and nationally) and how those changes are affecting people.

The CTS has two main components:

- Site visits to 12 communities
- Surveys
  - Household Survey
  - Physician Survey
  - Other: employer survey and survey of health plans (both discontinued)

Background Information on the Household Survey and the Physician Survey

Sample design:

- Complex sample design (e.g., stratification and oversampling)
- Most of the sample is drawn from 60 randomly selected *sites* (communities), of which 12 have larger samples and are therefore called the *high-intensity sites*. There is also a small *national supplement* to increase precision.
  - Same 60 sites in each survey.
  - The 12 high-intensity sites are:

Boston, Mass.	Miami, Fla.
Cleveland, Ohio	Northern New Jersey
Greenville, S.C.	Orange County, Calif.
Indianapolis, Ind.	Phoenix, Ariz.
Lansing, Mich.	Seattle, Wash.
Little Rock, Ark.	Syracuse, N.Y.

Types of estimates:

- Nationally representative estimates
- Site-level estimates for the 12 high-intensity sites. (Estimates for the other 48 sites are not generally considered reliable enough to be reported individually.)
- Estimates for people grouped by geographic characteristics
  - Type of state (e.g., states with "generous" Medicaid eligibility criteria)
  - Type of site (e.g., communities with high level of HMO penetration)

Public use and restricted use data files:

- Data files and documentation are available from the Health and Medical Care Archive (www.icpsr.umich.edu/HMCA) of the Inter-university Consortium for Political and Social Research (http://www.icpsr.umich.edu)
  - All *public use files* can be downloaded at no cost.
  - All *restricted use files* are available at no cost but require submitting an application to ICPSR to describe your plans for maintaining data confidentiality.
  - Documentation for both the public use and restricted use files can be downloaded at no cost (user's guides, codebooks).
- Documentation for the data files, as well as the survey methodology, is also available in the Technical Publications section of the HSC web site. (www.hschange.org/index.cgi?func=pubs&what=8/)
- Technical help is available at CTShelp@hschange.org.

ICPSR study numbers
(Go to www.icpsr.umich.edu and search by study number.)

Year of Survey	Household Survey Study Number	Physician Survey Study Number
1996-97	2524	2597
1998-99	3199	3267
2000-01	3764	3820
2003	[data files available Spring or Summer 2005]	[no survey]
2004	[no survey]	[data files available Spring or Summer 2006]

Visit CTSonline (www.hschange.org/ctsonline):

- CTSonline is a web-based system for generating user-specified tables showing results from the Household Survey and the Physician Survey.
- User specifies topic, year(s) of data, and how to disaggregate the results.
- National estimates only

CTS Household Survey

Years:

- 1996-97
- 1998-99
- 2000-01
- 2003

Sample:

- Represents civilian non-institutionalized population
- Total: about 30,000 families and 60,000 people
  - High-intensity sites: about 2000 – 2500 people per site
  - Low-intensity sites: average about 600 people per site
- There is some overlap of sample across rounds, but the data cannot be used for panel analysis.

Response rates: 59 – 65 percent

Survey content:

- Health insurance coverage
- Use of health services
- Unmet needs and expenses
- Usual source of care
- Patient trust and satisfaction
- Last visit to a medical provider
- Health status
- Chronic conditions
- Risk behaviors and smoking
- Employment, earnings, and income
- Demographic characteristics

Example of analysis using Household Survey data: Cunningham, Peter J., James D. Reschovsky, and Jack Hadley, *SCHIP, Medicaid Expansions Lead to Shifts in Children's Coverage,* Issue Brief No. 59, Center for Studying Health System Change, Washington, D.C. (December 2002). http://www.hschange.org/CONTENT/508/

- Findings include: Between 1996-97 and 2000-01, there was a decrease in the percentage of low-income children who were uninsured. This was the net effect of an increase in the percentage of low-income children with Medicaid/SCHIP coverage combined with a smaller decrease in the percentage of low-income children with private insurance.
- Examples of survey information used:
  - Family income as a percentage of poverty
  - Insurance coverage (private, Medicaid/SCHIP, other, uninsured)
  - State
- Use of geographic information: Part of this analysis groups people by whether the state where they live had a "large" expansion in eligibility for Medicaid and other state-run health insurance programs.

Technical issues:

- Site estimates (even for the 12 high-intensity sites) can have relatively large confidence intervals, so data users might want to pool observations from multiple years of the survey when calculating site estimates.
- Software for estimating standard errors
  - Specialized software is necessary because of complex survey design (i.e., not a simple random sample)
  - National estimates:
    - SUDAAN is optimal.
    - User's guides and HSC Technical Publication No. 40 explain when it might be reasonable to use other software packages (such as Stata and SAS).
  - Site-specific estimates: SUDAAN, Stata, and SAS all provide equivalent standard error estimates.

Public use and restricted use data files:

- Both files have the same observations, but the restricted version has more detailed information.
- The restricted version has:
  - Less editing of some variables
  - County identifiers (public version has only state and site identifiers)
  - Chronic conditions

CTS Physician Survey

Years:

- 1996-97
- 1998-99
- 2000-01
- 2004

Sample:

- Represents nonfederal patient-care physicians spending 20 or more hours per week in direct patient care
- Frame comes from American Medical Association (AMA) and American Osteopathic Association (AOA)
- Total: about 12,000 physicians each year
  - High-intensity sites: about 300 – 500 physicians per site
  - Low-intensity sites: average about 125 physicians per site
- More than half the sample overlaps across rounds, and those cases can be used for panel analysis.

Response rates: 59 – 65 percent

Survey content:

- Basic information on practice, specialty, and board certification
- Career satisfaction
- Physician time allocation
- Medical information obtained by patients
- Practice arrangements and ownership
- Priorities within practice
- Computer use
- Medical care management strategies and gatekeeping
- Scope of care
- Ability to provide care
- Ability to obtain needed services for patients
- Acceptance of new patients
- Practice revenue
- Compensation
- Race/ethnicity

Example of analysis using Physician Survey data: Cunningham, Peter J., *Mounting Pressures: Physicians Serving Medicaid Patients and the Uninsured, 1997 – 2001,* Tracking Report No. 6., Center for Studying Health System Change, Washington, D.C. (December 2002). http://www.hschange.org/CONTENT/505/

- Findings include: Between 1996-97 and 2000-01, there was a decrease in the percentage of physicians providing any charity care and a decrease in the percentage of physicians serving Medicaid patients.
- Examples of survey information used:
  - Hours of charity care provided in the last month
  - Percentage of practice revenue from various sources (managed care, Medicaid)
  - Acceptance of new patients by patient insurance type (private insurance, Medicare, Medicaid, uninsured)
- Use of geographic information: Table 3 has site-level estimates. The third column of numbers is the difference between 1996-97 and 2000-01. Note that site-level differences have to be pretty large in order to be statistically significant.

Technical issues:

- Counts of physicians are not reliable. Therefore, do not use the survey data to make such statements as, "There were [number] primary care physicians in Cleveland in 2003."
- Panel analysis: possible only for national estimates
- Site estimates (even for the 12 high-intensity sites) can have relatively large confidence intervals, so data users might want to pool observations from multiple years of the survey when calculating site estimates.
- Software for estimating standard errors
  - Specialized software is necessary because of complex survey design (i.e., not a simple random sample)
  - National estimates:
    - SUDAAN is optimal.
    - User's guides and HSC Technical Publication No. 40 explain when it might be reasonable to use other software packages (such as Stata and SAS).
  - Site-specific estimates: SUDAAN is optimal. Unfortunately, we have not had the resources to investigate cases when it might be reasonable to use other software packages for estimating standard errors for site estimates.

Public use and restricted use data files:

- Both files have the same observations, but the restricted version has more detailed information.
- Limitations of public version:
  - Public version cannot be used for site estimates, because there are no geographic identifiers on the file.
  - Public version is useful for preliminary analysis only, because there are no sample design variables on the data file, which means that the standard error estimates cannot be calculated.
  - Public version cannot be used for panel analysis.
- Restricted version has:
  - Less editing of some variables
  - Geographic identifiers (county, site, state)
  - Identifiers for panel sample
  - Sample design variables for use in estimating standard errors

Summary file:

- Data file with site means for selected Physician Survey measures for the 60 sites. It was created because the public use file has no site identifiers.
- Can be downloaded directly from the ICPSR web site
- Codebook lists estimates and standard errors for the 12 high-intensity sites