

Maintaining and Enhancing Representativeness of State Health Surveys

David Grant^{*}, Susan Scott[†], Nancy Breen[‡], J. Michael Brick^{**}, E. Richard Brown^{††}

Keywords: survey practice

DOI: [10.29115/SP-2011-0009](https://doi.org/10.29115/SP-2011-0009)

Survey Practice

Vol. 4, Issue 2, 2011

Maintaining and Enhancing Representativeness of State Health Surveys

Surveys providing local or small-population estimates are often administered by telephone but, in the last 20 years, obtaining survey responses has required increasing effort. For a number of reasons, U.S. survey telephone response rates have decreased from about 50 percent to about 20 percent. Low response rates undermine confidence that estimates accurately represent the study populations.

A November 2009 workshop, *Maintaining and Enhancing Representativeness of State Health Surveys: Lessons for the California Health Interview Survey (CHIS)*, was designed to discuss methods that CHIS is using to maintain response rates and evaluate sample representativeness. Organized by the Applied Research Program (ARP) within the Division of Cancer Control and Population Sciences (DCCPS) of the National Cancer Institute (NCI), the invited workshop attracted a broad range of researchers because these issues are central to all telephone surveys. Survey methodologists explored three pressing issues confronting CHIS and other random-digit-dial (RDD) telephone surveys:

- Estimates for local areas
- Ethnically diverse populations
- Nonresponse bias and coverage bias

The [California Health Interview Survey \(CHIS\)](#) is an RDD household telephone survey. CHIS is the largest population-based state health survey in the U.S. With its focus on health care access, health insurance coverage, health behaviors, chronic health problems, cancer screening, and other key health

* **Institution:** UCLA Center for Health Policy Research

† **Institution:** National Cancer Institute

‡ **Institution:** National Cancer Institute

** **Institution:** Westat

†† **Institution:** UCLA Center for Health Policy Research

issues, CHIS provides comprehensive data that serve a wide range of research and policy purposes. CHIS aims to produce high-quality, population-based data that are representative of California's geographic and demographic diversity. Therefore, the CHIS project team has an ongoing interest in meeting response rate, noncoverage, and related methodological challenges that potentially affect the representativeness of CHIS data.

ESTIMATES FOR LOCAL AREAS

As policymakers, program planners, and advocates call for more local data to inform decision-making, interest in estimates for small areas has intensified. Researchers also seek local data to identify where interventions are needed, design interventions, and monitor their implementation and effectiveness. Responding to these needs in California, CHIS provides estimates for the entire state and most counties, as well as sub-county estimates for the two most populous counties, Los Angeles and San Diego. The estimates enable counties to compare themselves with other counties and the state as a whole. CHIS frequently receives requests from policymakers and others for estimates of increasingly smaller or alternative geographic areas that are not easily accommodated by the CHIS geographically stratified sample design.

One option to generating health estimates for small geographic areas is to use statistical models such as small area estimation. Modeling supplements survey data from the target area/domain of interest with other data. CHIS produces direct estimates and to a lesser extent modeled estimates. Key questions that workshop participants raised were whether methods for obtaining small-area estimates could contain costs, address the methodological requirements of small geographic areas, and maintain data quality. While useful and sometimes very effective for generating estimates at the zip code level, for example, small area estimation is complicated, resource intensive, difficult to replicate, and potentially challenging for policy-makers and health advocates to embrace due to its complexity. For CHIS, small area estimation will likely continue as a useful, but limited, supplementary tool.

ETHNICALLY DIVERSE POPULATIONS

Collecting information from racial/ethnic and other small subpopulations is essential for addressing access to health care and other disparities. Collecting these data has multiple operational challenges, including inconsistencies in racial/ethnic categories across surveys and over time, inaccurate linguistic translation of survey instruments, and difficulties in hiring, training, and retaining qualified bilingual interviewers.

Workshop participants discussed several approaches to efficiently identifying members of rare groups for oversampling. One option is to stratify telephone exchanges linked to areas with a high density of a target ethnic group, but this method may be inefficient or biased if all strata are not sampled. Obtaining auxiliary lists from organizations with ethnic ties or ethnically associated

surnames is also possible, but this technique requires that the frame and the list share a common identifier and that the list be relatively complete, with broad coverage of persons in the rare group. Also, using government administrative data sets as auxiliary lists tends to pose a variety of data accessibility and quality challenges. A third option, network sampling, relies on the social networks of persons in an identified group; however, because the selection probability is typically unknown, this approach does not generate a probability sample of small or rare population groups.

CHIS combines surname list sampling and geographic oversampling of areas with high proportions of the target population to produce small-group estimates (i.e., Koreans and Vietnamese). After each data collection cycle, CHIS evaluates and modifies its sampling procedures for the next cycle to optimize the efficiency of oversampling rare groups. The methods used to oversample small population groups have been presented at survey methods conferences and are described in detail in the CHIS methodology reports (California Health Interview Survey 2008; Edwards et al. 2002).

COVERAGE AND NONRESPONSE BIAS

Surveys rely upon the representativeness of their samples to ensure that estimates describe the target population. Low response rates and poor coverage are the two major sources of bias that can undermine sample representativeness in RDD surveys. Studies of bias in CHIS data have found that coverage bias is a greater threat to their representativeness than nonresponse bias (California Health Interview Survey 2009; Lee et al. 2009).

COVERAGE BIAS

Coverage bias may occur when population members do not appear in the sample frame. A traditional RDD frame includes only households with landline telephones; because such samples exclude households that use cell phones, coverage bias may occur. Research from the National Center for Health Statistics (NCHS) has demonstrated that “cell phone only” households are likely to be younger, have lower income, lack health insurance, and report being in good or excellent health. CHIS has developed methodologies for including cell phone samples, piloting them in 2005 and implementing them in 2007 and 2009 (Brick, Edwards, and Lee 2007; Brick et al. 2010; Lee et al. 2010); however, cell samples are not suited well for small areas or rare groups.

An alternative to drawing a sample of telephone numbers is to use an address-based sample (ABS). The frame consists of mailing addresses of residential housing units within the geographic area of interest. A sample for each area can be drawn based on population density. This approach has challenges, again due to multilingual deployment and questionnaire complexity. Specifically, telephone numbers would be unavailable for about 40% of addresses and for almost all cell-only households. Switching to an ABS frame would require mail recruitment, which could introduce different errors

that are not well defined.

NONRESPONSE BIAS

Like other telephone surveys, response rates for CHIS are low and have declined significantly over time. Intensive follow-up analyses and benchmark comparisons suggest that little significant, systematic bias is present for sociodemographic variables, but no benchmarks exist to evaluate bias for most of the local-level health variables measured by CHIS. Consequently, CHIS has conducted several experiments to test methods for improving its response rates. Methods shown to improve response rates have been incorporated in the next survey administration, reported at survey method conferences, and described in the CHIS methodology reports (California Health Interview Survey 2008; Edwards and Brick 2006).

An encouraging finding is that small prepaid cash incentives improve response rates, or at least help to stem their decline, and CHIS has been using them since 2005. CHIS may benefit from additional research to understand the optimal dollar amount for an incentive and the impact of incentives at different stages of the interview process (screeener, extended, refusal). Workshop participants urged the CHIS project team to further explore incentive options with the Office of Management and Budget (OMB) and funders to build on previous findings (California Health Interview Survey 2008; Curtin, Singer, and Presser 2007; Singer, Hoewyk, and Maher 2000).

Mixed-mode administration approaches (e.g., mail survey followed by telephone interviews with nonresponders) have shown promise for improving response rates. However, it is unclear what kinds and how much bias may be introduced by changing from a single to a mixed-mode survey. Paper-based instruments do not permit the complex skip patterns possible with computer-assisted telephone interview (CATI)-administered surveys, nor are they conducive to multilanguage survey administration. Because information about mixed-mode data collection and bias is limited, workshop participants suggested embedding small-scale tests into CHIS data collection procedures to ascertain whether this approach can improve response rates.

EXPLORING NEW PATHWAYS TO STRENGTHEN AND SUSTAIN CHIS

As discussed, CHIS has tested ways to improve response rates and documented promising approaches. The CHIS pilot test of a cell-phone-only sample in 2005 addressed the feasibility of conducting an omnibus health survey by cell phone and explored methods to integrate cell phone and landline samples in the weighting process. CHIS 2007 included an area probability sample to systematically evaluate and quantify potential coverage bias and nonresponse bias.

Workshop participants identified methodological research needs in the following areas:

- Creating and testing alternatives to response rates to better measure representativeness
- Monitoring and evaluating alternatives to RDD sampling to sample small geographic areas
- Identifying the optimal combination of lists and geographic oversampling to produce representative estimates for specific ethnic subgroups
- Determining optimal incentives for reducing potential nonresponse bias
- Assessing the representativeness of survey data obtained through mixed-mode vs. single-mode RDD methodologies
- Exploring options to reduce coverage bias, including RDD with cell and landline samples, address-based sampling, and other/multiple frames
- Exploring cost trade-offs to maintain or increase response rates, such as whether the costs of intensive follow-up needed to modestly improve response rates are offset by savings in follow-up costs when incentives are paid to respondents

To read a full-length report on the November 2009 CHIS Workshop, visit http://appliedresearch.cancer.gov/surveys/chis/chis_methods_workshop2009.pdf

[Note: there are underscores in the hyperlink above after “chis” and “methods”.]

ACKNOWLEDGEMENTS

The authors would like to acknowledge other members of the workshop planning committee and other invited speakers. Planning Committee: Alyssa Grauman, National Cancer Institute; Sunghee Lee, University of Michigan; Brian Harris-Kojetin, Office of Management and Budget; Benmei Liu, National Cancer Institute; Van Parsons, National Center for Health Statistics, Centers for Disease Control and Prevention. Speakers: Mick Couper, University of Michigan; Robert Groves, U.S. Census Bureau; Karol Krotki, RTI International; Stephen Immerwahr, New York City Department of Health and Mental Hygiene; James Jackson, University of Michigan; Nancy Mathiowetz, University of Wisconsin-Milwaukee; David Takeuchi, University of Washington. We wish to acknowledge Caroline McLeod, NOVA Research Company for outstanding documentation of the meeting.

REFERENCES

- Brick, J.M., W.S. Edwards, and S. Lee. 2007. "Sampling Telephone Numbers and Adults, Interview Length, and Weighting in the California Health Interview Survey Cell Phone Pilot Study." *Public Opin Q* 71: 793–813.
- Brick, J.M., I. Flores-Cervantes, S. Lee, and G. Norman. 2010. "Nonsampling Errors in Dual-Frame Telephone Surveys." *Surv Methodol* 37 (1): 1–12.
- California Health Interview Survey. 2008. "CHIS Survey Design and Methods: Methodology Reports." <http://www.chis.ucla.edu/methodology.html>.
- . 2009. "CHIS 2007 Methodology Series Report 2: Data Collection Methods." Los Angeles, CA: UCLA Center for Health Policy Research.
- Curtin, R., E. Singer, and S. Presser. 2007. "Incentives in RDD Telephone Surveys: A Replication and Extension." *J Off Stat* 23 (1): 91–105.
- Edwards, W.S., and J.M. Brick. 2006. "Effects on Response Rates of Multiple Sponsors on Advance Letters for an RDD Survey." Presented at the Annual Meeting of the American Association for Public Opinion Research, Montreal, Quebec, Canada.
- Edwards, W.S., J.M. Brick, I. Flores-Cervantes, C. DiSogra, and W. Yen. 2002. "Sampling Race and Ethnic Groups in RDD Surveys." *Proc Surv Res Methods Sec Am Stat Assoc*, 893–98.
- Lee, S., J.M. Brick, E.R. Brown, and D. Grant. 2010. "Growing Cell-Phone Population and Noncoverage Bias in Traditional Random Digit Dial Telephone Health Surveys." *Health Serv Res* 45 (4): 1121–39.
- Lee, S., E.R. Brown, D. Grant, T.R. Belin, and J.M. Brick. 2009. "Exploring Nonresponse Bias in a Health Survey Using Neighborhood Characteristics." *Am J Public Health* 99 (10): 1811–17.
- Singer, E., J. Van Hoewyk, and M.P. Maher. 2000. "Experiments with Incentives in Telephone Surveys." *Public Opin Q* 64 (2): 171–88.