THE REACH 2010 COMMUNITY-BASED SURVEYS AND THEIR COST ISSUES

Rachel M. Harter and Carol-Ann Emmons
NORC, 55 E. Monroe, Suite 4800, Chicago, Illinois 60603

KEY WORDS: Community Survey, Screening, Random Digit Dial, Coverage

Introduction

Community-based surveys pose special challenges not faced by national or statewide surveys. For example, the community itself may not be defined in a way that lends itself to statistical sampling methods. Also, traditional RDD sampling methods can be less efficient on a smaller geographical scale. Third, idiosyncrasies in the target population make it risky to build the sample design on assumptions from experiences elsewhere. Finally, local attitudes and sensitivities regarding participating in surveys affect decisions about how to approach potential respondents. All of these issues and more were faced by NORC in designing and collecting the first round of data for REACH 2010.

This paper first describes the REACH 2010 program and its surveys. Particular REACH communities are used as examples to illustrate the problems commonly faced by community surveys. Because REACH involved fielding 21 separate community surveys, simultaneously, the problems I will describe were compounded, leading to significantly higher-than-budgeted costs. This in turn prompted NORC to re-evaluate many aspects of the survey. Some of the initiatives resulting from this re-evaluation, and their implications will be described in greater detail by other authors in this session. This paper will provide an overview of the issues, investigations, and solutions that led to a much more cost-efficient survey process for the REACH communities.

The REACH 2010 Program

Racial and Ethnic Approaches to Community Health (REACH) 2010 is the flagship program of the Center for Disease Control & Prevention (CDC) to eliminate racial and ethnic disparities in health through demonstration projects in selected communities around the country (http://www.cdc.gov/reach2010/). Each community received a grant to design and implement a health intervention involving one or more ethic or racial minority groups, and targeting one or more health priority areas. The minority groups are African American, Hispanic, Asian, Native American, and Pacific Islander. The health priority areas are: diabetes, cardiovascular disease, breast and cervical cancer, HIV/AIDS, infant mortality, and adult and childhood immunization.

The health interventions are designed to increase community members' knowledge and awareness of healthy lifestyle choices, as well as improve access to medical care. For example, in the Rio Grande Valley in Texas, diabetes strikes nearly 25% of the Hispanic/Latino population. The REACH *Promotora* Community Coalition there, led by the Migrant Health Promotion organization, designed an intervention that consisted of three parts: 1) a school-based primary prevention program to educate youngsters about the health benefits of eating right and getting exercise; 2) a clinic-based secondary prevention program focused on proper self-care, and 3) a *colonia*-based (*colonias* are rural, low-income neighborhoods) program in which trained *promotoras* (peer educators) offer in-home support to diabetics and their families focusing on nutrition, physical activity and regular health care, as well as leading walking groups and cooking classes with *colonia* residents. In other communities such as San Diego, the community coalition is training beauticians to be peer-educators about diabetes with their African American clients, teaching cooking classes with healthier adaptations of traditional foods, and encouraging proper self-care for the prevention and treatment of diabetes. In Chicago and Las Vegas, influential religious congregations in the African-American communities sponsor health screenings and healthier lifestyle choices. In San Francisco, the community coalition is working with Vietnamese women to prevent breast and cervical cancer by training lay educators to do community outreach, establishing a clinic staffed by a Vietnamese physician to provide pap tests to low-income, uninsured Vietnamese women, and setting up a pap smear registry and reminder system for Vietnamese women in Santa Clara County.

Altogether, the CDC has given grants to more than 30 communities around the country, and has contracted with NORC to collect data in the communities shown in Figure 1 (21 communities the first year, 27 communities in subsequent years). NORC is collecting survey data in these communities over several years so that the CDC and the communities can monitor progress in the effectiveness of the health interventions. Data collection is to be completed by telephone, if at all possible, provided the community has at least 80 percent telephone coverage. Although the same survey instrument, which is based on the CDC's Behavioral Risk Factor Surveillance System (BRFSS; Centers for Disease Control & Prevention, 2003), is used across all communities, the sample designs must be

tailored to each community. Thus we have, in effect, 27 separate community-based surveys.

The Challenges

One of the biggest challenges faced by NORC was obtaining a definition of each community that permitted probability-based sampling. For REACH, each community was defined in terms of two dimensions--geography and race/ethnicity.

The racial/ethnic definitions of the communities were not always straightforward. For example, we negotiated with one Native American community over whether survey eligibility should be based on simple self-report, or possession of a Certification of the Degree of Indian Blood (CDIB) card. We also debated the inclusion of various nationalities in Asian communities. For example, in one community, focused on Cambodians, the question arose whether persons of Cambodian descent who migrated to American from Vietnam should be considered Cambodian or Vietnamese. One or two communities had to define for themselves whether recent African immigrants are part of the same community of African-Americans as the American slave descendants.

Geographically, the communities varied considerably in size from a few census tracts or a few zip codes, to city or county-level boundaries, to the entire state of Oklahoma. Congregation-based communities were particularly challenging to define because the churches were understandably unwilling to provide us with membership lists, and their catchment areas were geographically large, and sometimes of low density with respect to the minority group of interest.

Another challenge with respect to defining the targeted geography is whether respondents can accurately say whether or not they live within the boundaries of a certain geography. Again, the most difficult geographical areas were defined by tracts, or by some special designation. The Atlanta Empowerment Zone, for example, is a well-defined urban area targeted for redevelopment, but most residents of Atlanta have no idea what its boundaries are or whether they live within the zone. In such cases we had to redefine the target geography in terms of zip codes or well-known street boundaries.

Once defined, some of the geographical areas did not lend themselves well to telephone surveys. The small communities defined by tracts or zip codes, in particular, do not coincide well with the coverage areas of telephone exchanges, leading to difficult tradeoffs between cost efficiency and coverage.

Ethnic considerations sometimes influenced mode of data collection. For example, Native Americans in Oklahoma were thought to have less than 80 percent telephone coverage, based on 1990 Census data, so we initially planned an area probability design with in-person interviewing in Oklahoma. However, the leadership of certain tribes objected to interviewer presence on tribal land (even though the interviewers were to be local recruits), so we ultimately reverted to a telephone survey despite the telephone coverage issue.

Characteristics of the target population also influenced the sample design. These included the density of the racial/ethnic group as a proportion of the total population within the geography; whether there existed pockets of relatively higher concentrations of the racial/ethic group so that stratification was a reasonable approach; and whether the group was so rare and thinly dispersed throughout a large geographical area that random digit dialing (RDD) was cost prohibitive. The last situation led us to consider the use of targeted surname lists as a cost-efficient alternative, or at least a supplement, to RDD.

When planning a design for a survey, the statisticians must rely on certain assumptions about the vendor's telephone sample as well as assumptions about the population itself. On the vendor side, we need to assume certain coverage rates, working residential number rates, and geographical hit rates. While our assumptions at the national level are quite good based on considerable experience with the vendor, our assumptions for these small communities are largely untested. For example, lacking data on the working residential number (WRN) rate for random digit dial (RDD) samples at the community level, we initially assumed a rate of 50 percent, based on our experience with national samples from the vendor we used to generate the RDD samples. What we found, however, is that the WRN rate actually varied considerably by community. Part of the variance was explained by the vendor's rotating maintenance schedule, as different parts of the country are updated at different times. The remainder we assumed to be true local variation. Similarly on the population side, while national census figures for large racial ethnic groups were available, we sometimes had to rely on 1990 PUMS data at the planning stage because 2000 census data were not yet available and national survey data were insufficiently detailed for such small groups.

Thus NORC went into these surveys with more than the usual amount of uncertainty about our design assumptions. We coped with the high level of uncertainty by using sample replicates. The replicates, therefore, functioned as far more than a phone shop management tool; replicates allowed us to make adjustments to our design at interim points in the data collection process.

As mentioned previously, the numerous challenges for sampling at the community level, compounded by the number of communities with which we were working simultaneously, resulted in significantly higher costs which required immediate

attention. NORC pulled out all the stops to address the problem. With the CDC fully aware of the challenges we faced, we formed a company-wide team to examine every aspect of the study. Our investigations can be categorized into the following four areas: 1) sample design, 2) improving response rates, 3) operational efficiencies, and 4) scheduling/planning/budgeting. Some specific initiatives and their implications will be discussed by other authors in this session. Here we give an overview of the nature of the investigations.

Sample Design

In the initial design stage, we spent considerable effort tailoring each community's design to the characteristics of the racial/ethnic group of interest within the targeted geography. These characteristics included the group's relative density and distribution within the target geography, estimated telephone coverage, assumed willingness to participate in a survey conducted over the telephone, and language needs. We considered cost implications as well as statistical implications. When we were faced with severe cost overruns, we revisited the tradeoffs and revised some of the decisions in favor of reducing costs at the expense of statistical issues. In other words, our revised designs sometimes led to sacrifices in effective sample sizes, either through increased clustering or increased variability in weights, for the sake of cost effectiveness.

We achieved the most dramatic increase in cost efficiency by adding supplemental samples of listed telephone numbers to our RDD samples. Because a higher percentage of such numbers are in fact working and linked with residential households, this change alone dramatically reduced the amount of sample phone numbers needed and the number of calls needed to complete screeners. However, when the listed phone numbers were added to the first round samples, the listed and RDD frames overlapped, complicating the probabilities of selection and the weights. This issue is discussed by Pedlow and Porras (2003). In the second round of data collection, we separated the listed telephone numbers from the numbers not known to be listed into formal strata in order to simplify the construction of weights.

A second change we made was to relax the requirement that half of the interviews in each community be with racially eligible females age 40-64. The problem was that the oversampling of females caused us to subsample out other eligible adults, thus increasing the initial sample size. Although we continued to oversample females, relaxing this requirement enabled us to complete the total number of interviews required with fewer households.

A third, and closely-related design change involved the number of eligible adults selected from each eligible household. Because of the difficulty of finding enough females in the 40-64 age group, we had allowed for multiple respondents per household from the start. However, we had restricted the number of other adults to one per household. We loosened that restriction somewhat, allowing for the selection of up to two other adults per household, thereby reducing the number of households needing to be screened.

All three of the changes described above were implemented by replicate so that weights could be calculated properly.

Response Rates

Any time we can improve response rates without increasing the number of calls we make, we save money. Reducing the number of calls saves even more money. With this in mind, we re-evaluated aspects of the survey that we felt impacted response rates.

We spent considerable effort fine-tuning the survey introduction and the advance letter. While many households will hang up before the interviewer can complete a single sentence, if you can get the respondent through the introduction, you have a much greater chance of completing at least the screening interview. We found that it was especially important to revise the introduction after the anthrax scares in the fall of 2001. Respondents who may not have been aware of the Centers for Disease Control were now frightened by an introduction that mentioned the CDC and health problems in the community. Therefore, we changed the survey introduction to mention the University of Chicago rather than the CDC.

The increase in listed telephone numbers in turn increased the number of households for which we had address information. This allowed us to send an advance letter to a higher proportion of the households in our sample. We also conducted tests with three different address-matching vendors to identify one with the highest match rate and degree of accuracy [Murphy, et al. (2003)]. This further increased the proportion of respondents to whom we were able to send an advance letter. Third, we adjusted the timing of the mailing of the advance letter so that households would receive the letter just a few days before we called them on the telephone. Finally, we revised the advance letter we had been using by shortening it and by placing greater emphasis on the health benefits to the community of survey participation. In year 2 of data collection, we also personalized the letter by addressing it to a specific household member. Anecdotal information from interviewing staff suggests that this change has caused more respondents to spontaneously mention the advance letter when the interviewer calls.

Our quest to improve response rates also led us to analyze the screening interview data we had collected to date to determine the point in the interview where we were most likely to lose respondent cooperation. We

found that a large percentage of interviews were lost at the geography screening questions. The questions that asked whether a respondent lived between certain street boundary were most problematic. Many respondents did not know north, south, east, and west, or even if they lived between two major roads. This problem led us to consider geocoding respondent addresses to determine in advance whether the addresses were within the targeted geography. This eliminated the need to ask the geography questions for significant portions of the samples, thereby improving the response rate for the screening interview.

Operational Efficiencies

Similar to improving response rates, operational efficiencies that allow us to complete the same number of interviews with less effort will save money. To achieve more operational efficiency, we concentrated our efforts on the calling rules we were using. We were required by contract to make up to seven calls to each telephone number, at different times of day, over a period of two weeks to complete a screening interview. We analyzed the productivity of additional calls over the required seven to determine a more efficient calling strategy that would not sacrifice much in terms of response rates. Harter et al. (2003) discuss this in greater detail.

We also reviewed interviewing production by shift to identify the most productive days and time periods for completing interviews. We then adjusted our staffing to have more interviewers working at times found to be most productive, and fewer interviewers working during less productive times.

We have found that for a large, ongoing survey like REACH, maintaining a steady interviewing workforce with a relatively constant work load results in greater efficiency (Buckley et al., 1998). Therefore, we endeavor to keep the workflow relatively constant year round, overlapping the close down of one phase with the start up of the next.

Finally, we analyzed the relationship between the productivity of the sample and its age (i.e., the time between when the sample was ordered and when it was worked in the telephone center). Not surprisingly, we found that the older the sample, the less productive it was in terms of generating completed interviews. The main reason for this is that more numbers become disconnected. More disconnected numbers leads to more sample being needed, more calls being placed, and greater discouragement among the interviewers. For many "quick-turnaround" surveys, the age of the sample is not an issue. For a large, ongoing set of surveys such as REACH, the age of the sample can have a significant impact on the level of effort required. Therefore, we tightened our scheduling process to minimize the time between when a batch of sample was ordered and when it would be worked in the telephone center, referring to our approach as "just in time" sampling.

Scheduling/Planning/Budgeting

We have already discussed several areas in which we tightened our control of the data collection schedule in order to achieve greater operational efficiencies. These were the mailing of the advance letter, the scheduling of interviewing staff, and ordering and releasing of fresh sample. We also developed processes and tools to provide ourselves with more timely and more accurate information that allowed us to further tighten our control of the data collection schedule. These included better sample monitoring reports, more frequent and direct communication among all the task leaders, including the statisticians and telephone shop management, and tools for better tracking our actual costs and managing future spending. These gains set a model that is being copied and implemented throughout the organization for other projects.

Conclusion

As a result of all the analyses and adjustments that were made we succeeded in reducing the interviewing time by 65 percent, and our interviewing costs by $1.3M. We also developed tools and standards that will benefit future surveys.

As mentioned previously, some of our design-based cost cutting measures came at the expense of effective sample size. During the second round of data collection, our cost-cutting measures were more successful than budgeted, leading us to reconsider some of the design-based decisions. As we continue our partnership with the CDC, we will continually revisit these issues until we find the right balance between cost containment and statistical efficiency.

While this discussion of cost issues might be viewed as an embarrassment because of the challenges we faced, we have chosen to highlight the issues in the interest of scholarly exchange and advancement of the industry as a whole. This discussion is intended to serve as a reminder that we in the survey industry need to be vigilant about all aspects of our work. Furthermore, we statisticians have a unique contribution to make in the success of our organizations and in our industry's ability to promote the public welfare through well-designed and efficiently-executed surveys.

References

Buckley, P. , J.M. Dennis, C. Saulsberry, V. G. Coronado, T. Ezzati-Rice, E. Maes, A. Rodin, R.A. Wright (1998), "Managing 78 Simultaneous RDD Samples," *1998 Proceedings of the American Statistical Association,* Survey Research Methods Section,

American Statistical Association, pp. 957-961.

Centers for Disease Control & Prevention (2003), "Health Risks in the United States: Behavioral Risk Factor Surveillance System," *At A Glance Annual Report*, Atlanta, GA.

Harter, R., Emmons, C., O'Muircheartaigh, C., and Murphy, W. (2003), "Optimizing Call Strategies in RDD: Differential Nonresponse Bias and Costs in REACH 2010," *2003 Proceedings of the American Statistical Association*, Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.

Murphy, J., Murphy, W., and Burke, A. (2003), "Assessing the Validity of Matched Address Data for Unlisted Telephone Numbers in REACH 2010," *2003 Proceedings of the American Statistical Association*, Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.

Pedlow, S. and Porras, J. (2003), "Outlier Weight Issues in REACH 2010," *2003 Proceedings of the American Statistical Association*, Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.

Figure 1.
REACH 2010 Survey Sites, 2001-2002