

Design and Operational Changes for the REACH U.S. Risk Factor Survey

Whitney Murphy, Rachel Harter, and Kanru Xia

National Opinion Research Center, 55 E. Monroe St., Chicago, IL 60603

Abstract

Increasing use of cell telephones, and corresponding decrease in households with landline phones, has led to coverage concerns in RDD surveys. For the Racial and Ethnic Approaches to Community Health (REACH) U.S. Risk Factor Survey, a set of CDC-sponsored community surveys, NORC interviews minority households, for which the landline coverage issue may be more acute. Recently, the REACH U.S. Survey converted to an address-based sampling approach with multi-mode data collection. Applications of existing methodologies were challenging due to limited availability of paradata to guide design and operational planning. In this paper we discuss the address-based design, alternatives considered, and related assumptions regarding response rates, eligibility rates, and other factors affecting sample size calculations and costs. We also discuss design and operational adaptations made as we gained experience.

Key Words: address-based sampling, household survey, multi-mode data collection, subpopulations

1. Introduction

In many U.S. communities, minority populations suffer disease problems disproportionately to the majority population. The particular diseases of concern are diabetes, cardiovascular disease, breast and cervical cancer, and HIV/AIDS. For example, the Hispanic population in southern Texas suffers diabetes in epidemic proportions. As another example, Asian women in some communities are less likely to get regular screening for breast and cervical cancer, resulting in disease occurrences being caught at later stages less responsive to treatment. The issues are complex. However, some of the underlying issues, including access to care, health education, informed self-care options, and health related risk factors, can be addressed with the participation of the minority populations. In 1999 the Centers for Disease Control and Prevention (CDC) inaugurated the Racial and Ethnic Approaches to Community Health program, initially called REACH 2010, to address these issues over the 2000 to 2010 decade. The CDC made cooperative agreements with 40 community coalitions and funded community interventions to address the local issues directly with the minority populations. The goal was to decrease or eliminate health disparities in the minority populations. Over time, the program became known as REACH U.S. to indicate its planned continuation beyond the initial decade. See <http://www.cdc.gov/reach/index.htm> for more information about the REACH program.

The CDC contracted with a survey research organization to collect data annually among the minority populations in around 28 of the communities with the interventions. Many of the survey questions are comparable to those of the Behavioral Risk Factor Surveillance System (<http://www.cdc.gov/BRFSS>). By collecting these data annually, the CDC and

the communities can track healthy lifestyle behavior over time to see whether the interventions are having an impact. The National Opinion Research Center at the University of Chicago (NORC) has been the data collection contractor for most years of the program.

The design and methodology of the REACH surveys have undergone many changes over the years. Each community has its own unique issues, requiring some tailoring as we gain experience with them. On a larger scale, the survey industry has faced challenges to Random Digit Dial (RDD) methodology as more and more households utilize cell telephones exclusively or nearly so. NORC transitioned to address-based sampling for REACH U.S. to compensate for reduced coverage of traditional RDD surveys, and the new methodologies also required tweaking to balance competing needs for coverage, timeliness, and cost containment.

In the following sections, we discuss many of the adaptations NORC made to the REACH 2010 and the REACH U.S. surveys. We begin with adaptations for community-specific design issues. Then we discuss the conversion to the address-based sampling design and subsequent revisions.

2. REACH 2010 Survey Design Adjustments

Prior to the REACH surveys, NORC had extensive experience with national surveys and national oversamples of some minority groups. NORC also had experience with surveys in local communities, sometimes with subsets of the population. However, the targeted nature of these surveys meant that nobody had specific knowledge about these communities for optimal designs. We had to make our best guesses, and adapt as we went along, to achieve 900 completed interviews per community per year, including minimum requirements by age range and gender. In this section we discuss revisions to mode, screener questions, rate assumptions, persons per household, and frame source. These changes were previously discussed in greater detail by Harter and Emmons (2003) and Murphy et al. (2003a, 2003b).

When the REACH 2010 survey was initiated, computer assisted telephone interviewing (CATI) with RDD samples was the preferred mode of data collection. In-person data collection was preferred in some communities where telephone penetration was lower or where cultural modesty suggested in-person interviewing may be more successful in gaining cooperation. Even in these matters, we sometimes found that our initial assumptions were wrong, and we changed modes in subsequent annual cycles as appropriate.

The screening questions underwent revision over time. One of the biggest challenges for the RDD screener was verifying that the household resided within the targeted geography. In some cases it was difficult for the communities to articulate their geographical boundaries, such as the area corresponding to congregation-based intervention programs. In other communities the geographies were extremely precise, but difficult to articulate in a way the respondents could easily understand and confirm their eligibility. For example, very few households know their census tract designation. They may not know if they are east or west of a specific major highway. And the boundaries of the Atlanta Empowerment Zone are not well known to residents of Atlanta. In addition to modifying the screener, in some communities we implemented a procedure for geocoding the addresses corresponding to the telephone numbers, where addresses could be

matched, to determine geographical eligibility and eliminate the need to ask the respondents. Even the introductory paragraphs to the screener were revised. After the 2001 anthrax scare, we found that our introduction stating CDC's concern about disease problems in the community alarmed people unnecessarily and required revision.

It was not unusual for our assumed rates of eligible households, cooperation, and completion to be somewhat off. It is common for participation rates to vary geographically and by racial/ethnic group, and we adjusted rates as we gained experience in these communities. For surveys of this nature, it is important to consider household size and composition of the minority population in addition to population size. We must also allow for migration and fluctuations since the latest decennial census. (The American Community Survey was not yet available for REACH 2010 at the program's inception.)

For some communities, minor adjustments to the rates were insufficient for us to reach our goals for completed interviews within our allotted budget. More substantial design changes were warranted, especially to identify eligible persons more efficiently. While REACH always permitted more than one respondent per household, we adjusted our rules for taking persons in the rarer age/gender categories at a higher rate in all communities.

To identify Hispanic and Asian groups more efficiently, we supplemented RDD with targeted surname lists. For various racial/ethnic/nationality groups, we stratified the communities by high and low density geographic areas and sampled at a higher rate in the high density stratum. While some of these design features were planned from the start in a limited number of communities, we increased use of such methods while monitoring the resulting design effects.

3. Coverage Concerns with Cell Phones

The changes described for REACH 2010 are not unusual for surveys of specific subpopulations. The REACH 2010 survey also faced national trends in cell telephone usage that increased the risk for coverage bias. According to the National Health Interview Survey, during the first half of 2009, 25 percent of American homes had only cell phones. In addition, 15 percent received most or all of their calls on cell phones even though they had a landline phone (Blumberg and Luke, 2010). This sums to 40 percent of American homes that are not reachable by traditional RDD.

Moreover, some minority populations have embraced cell phones at a higher rate than the majority, increasing the risk for coverage bias in REACH communities. Hispanic adults, one of the REACH target populations, are more likely to live in cell phone only households than the other racial/ethnic groups. Adults with the following characteristics are more likely to have only cell phones: those who are young (aged 18-34), male, renting, living alone or with non-relatives, living in central cities, living in multi-unit buildings, unmarried, or poor (Blumberg and Luke, 2010; Tucker et al., 2007).

Besides demographic differences between cell phone only and landline populations, previous studies have also found differences in health estimates between the two groups. Adults living in cell phone only households were found to be less likely to have health insurance coverage or exercise regularly and were found to drink and smoke more;

however, these studies also found they had less high blood pressure and diabetes mostly due to their young ages (Link et al., 2007). All of these made us concerned about the potential coverage bias of traditional RDD and pushed us to adopt what has become known as address-based sampling (ABS). The REACH U.S. Risk Factor Survey was one of the first major surveys to adopt ABS.

The REACH U.S. Survey is conducted in phases, with each annual cycle of data collection as a phase. As of this writing, NORC is in the late stages of data collection for Phase 2. In the following section, we describe the ABS for Phase 1, and the modifications to the design that were necessary to achieve the performance goals. Then we discuss the revised design for Phase 2.

4. REACH U.S. Survey Phase 1 Address Based Design

The REACH U.S. Survey is very similar to the REACH 2010 Survey in scope and purpose. NORC's role is to collect 900 interviews in each of 28 communities (Figure 1), some of which are identical or nearly identical to a REACH 2010 community. As in REACH 2010, a few communities have health focuses that introduce an additional requirement that 300 of the completed interviews be with household members who are either females age 40 to 64, or adults age 65 and older. The surveys for the two contracts are very similar, and the screening method of household rostering and person selection is nearly identical. However, REACH U.S. made a significant design change from RDD to ABS.

Address-based sampling shifts the primary sampling unit from the telephone number—as used in landline RDD sampling—to the mailing address. This is expected to increase coverage, through the use of a frame that is expected to contain over 95 percent of the households in the U.S., while still allowing for the use of traditional data collection methods such as telephone, mail, and in-person interviewing. For REACH U.S., the frames for each community area were based on the United States Postal Service's Delivery Sequence File (DSF), which contains addresses for nearly all of the households in the country. This file was narrowed to contain only addresses within the targeted geography for each community, thereby resulting in a sampling frame for each community with near complete coverage of households. From these frames, random samples for each area were selected.

Phase 1 of REACH U.S. employed a three-pronged data collection approach. First, the sampled addresses were matched to telephone numbers. Telephone interviewing is often the most cost-effective approach, and telephone matches were expected to be high, so it was assumed that most interviews would be collected via telephone. Any household without a telephone match was initially sent a postcard in order to pre-screen households for eligibility as well as obtain a phone number. Self-administered questionnaires (SAQs) were mailed to all households without a phone match, and later to those who did not respond via telephone. Finally, an in-person follow-up to a subset of non-responders was conducted. This was expected to help boost response rates and cover hard-to-reach households, thereby reducing some potential non-response bias.

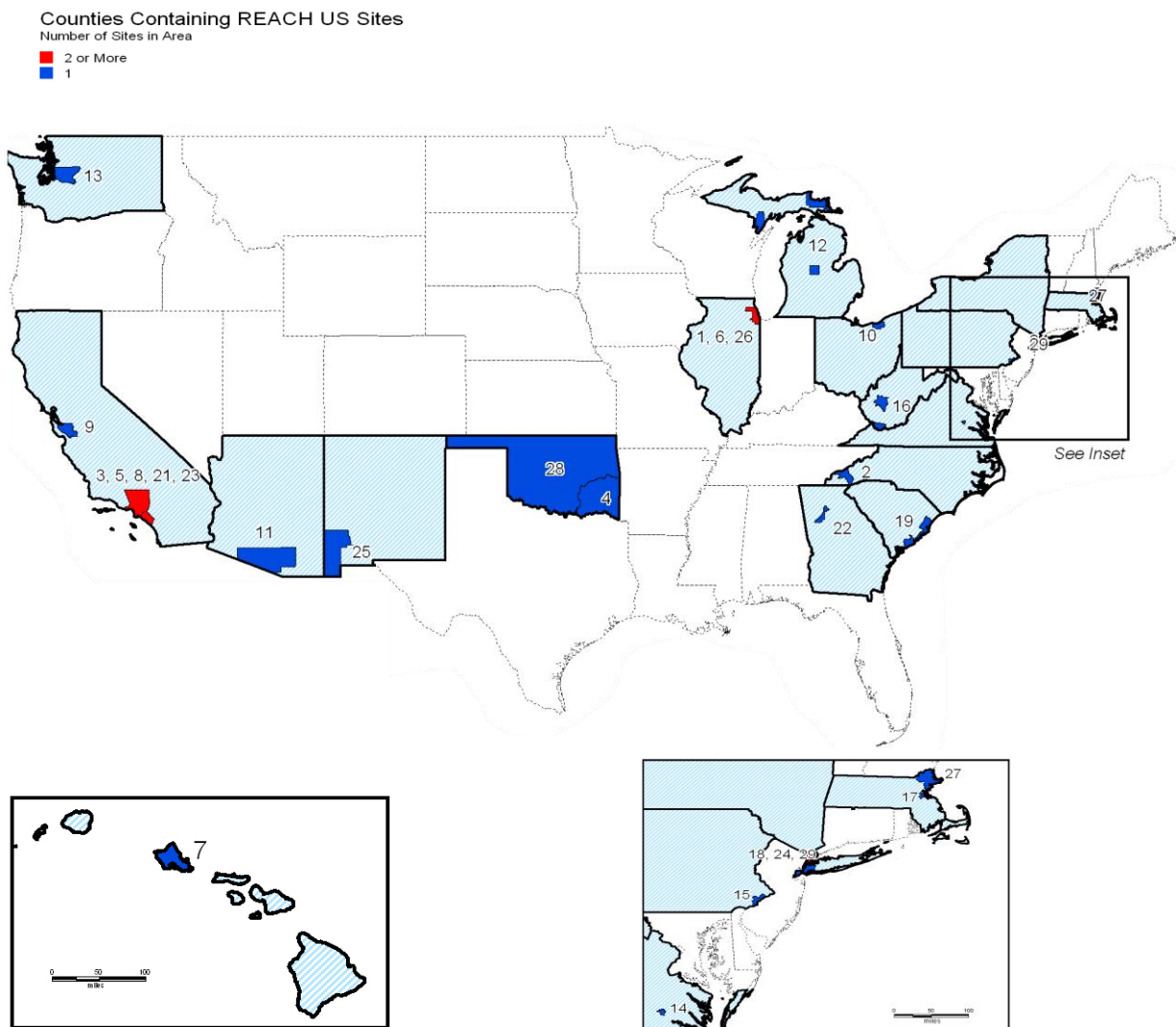


Figure 1: REACH U.S. Survey Communities (dark blue and red shaded areas)

Because this approach was relatively new and not attempted before on this large a scale, or in these rare and specific geographic and racial/ethnic populations, initial assumptions were based on RDD and mail surveys and on the few address-based studies with published results (Link et al., 2008). In RDD surveys, we typically found addresses for approximately 60 percent of telephone numbers; for ABS, we expected about the same for the reverse matching (address to telephone number) process, plus more from the mailed screeners. We also assumed that, of the households for which we obtained a telephone number, most (around 90 percent) would tell us that the telephone number matched the address correctly (i.e., that they lived at the household we had sampled). Rate assumptions for eligibility, screener completion, and interview completion for the postcard, telephone, SAQ, and in-person follow-up, were based on previous similar studies, including REACH 2010.

After several weeks of data collection, it was clear that many of the assumptions we had made did not hold for this ABS study. First, the match rate, while close to 60 percent in many communities, was as low as 35 percent in other communities. Furthermore, apartments, which made up a large proportion of many of our communities' housing units, proved much more difficult to match than single family units. However, instead of receiving fewer matches for apartments, we often received the same telephone number matched to several apartments within the same building in our sample, thus not only requiring a detailed de-duplication process, but resulting in the loss of many initial presumed matches.

Second, and on a related note, the address confirmation rate was much lower than expected, and was especially low in communities with a large percentage of apartments. Due to this phenomenon of receiving the same phone number matched to several sampled addresses, it is not surprising that many screener respondents contacted via telephone did not confirm their address to be the one in our sample. While many of these are expected to be due to a bad match, it is also likely that some are hidden refusals. It should be noted, too, that other ABS studies in the literature did not confirm the addresses during the telephone interviews, accounting for some discrepancies in yield rates.

Finally, the postcard return rate was abysmal. Other studies on which we based our assumptions had a relatively decent response with a non-negligible number of households sending back a phone number. REACH U.S. experienced only a two percent response to the postcard. The postcard was abandoned after the first replicates of sample due to its failure. Its failure also impacted the confidence in and the use of SAQ mailings in Phase 1. The original design was to mail SAQs to all households without a phone match, as well as to all households for which the address was not confirmed on the phone, or for which a phone interview was otherwise not completed. Instead, SAQ mailings were discontinued for all replicates after the first set, thus relying almost completely on telephone in Phase 1, and thereby reducing the ability to make valid SAQ response assumptions for Phase 2. (However, based on the first replicates, the SAQ did reasonably well in most communities in Phase 1.) On a positive note, screener, interview, and eligibility rates were close to assumptions.

Due to these worse-than-expected rates in Phase 1, and the resulting budget concerns, we revised our methodology mid-year based on revised expectations. First and foremost, we made revisions to the frames from which our samples were drawn to make them as efficient from the outset as possible, while still maintaining statistical integrity. We did this by instituting a dual-frame design, integrating race-targeted list-based addresses into the ABS frame to enhance its efficiency. Our vendor provided this race-targeting/identification service, which was expected to cover approximately half of the target population but have a hit rate¹ of 60 percent or higher. We actually found the hit rate to be closer to 80 percent in most communities. This targeting is readily available only for some racial/ethnic groups (e.g., Hispanics and Vietnamese), but is less effective and sometimes impossible for other groups (e.g., African Americans and some rare Asian subgroups). We created mutually exclusive "listed" and "unlisted" portions of the frames from which we drew the remainder of our samples. The new sample was expected to be

¹ The hit rate is defined as the percentage of the targeted addresses that were expected to contain a household member in the targeted group(s). For REACH, we expected at least 60 percent of the race-targeted addresses actually to be households with members in the targeted racial/ethnic group(s).

much more efficient than the original ABS sample. We also supplemented with some RDD sample in a few of the most challenging communities at the end of data collection in order to complete the remaining interviews before our deadline. Naturally the false starts and revisions came at a price. While the in-person follow-up component of data collection was intended to improve response rates and alleviate some potential bias, the remaining budget was insufficient to collect enough in-person interviews to have much impact.

The many changes to the REACH design required careful consideration of the weighting methods. The REACH 2010 weights had followed the usual sequence of base weights based on probabilities of selection with adjustments for eligibility and nonresponse at the screener and interview stages. The probabilities of selection for REACH U.S. were more complex, given the multiple frames. Ultimately the base weights were streamlined and poststratified to marginal totals from the American Community Survey, where available, or to projected census figures.

5. REACH U.S. Survey Design Adjustments for Phase 2

In Phase 2 of REACH U.S., several design and operational changes were made. First, a dual-frame design was used from the start, with mutually exclusive listed (race- or race- and age-targeted) and unlisted frames constructed for each community. Samples were drawn disproportionately from these frames, with more (usually significantly more) being drawn from the listed frames. To boost efficiency and ensure a rich enough frame in some of the most difficult communities, we constructed a third frame of eligible households identified in Phase 1. While this was not desirable from the standpoint of obtaining a unique cross-sectional sample each year, it was sometimes necessary due to the small size of the frame and the difficulty in identifying the eligible households during Phase 1. Finally, because of the disproportionate expense of in-person interviewing, we dropped this mode in Phase 2 for most communities. However, one community was interviewed exclusively in person in Phase 2.² One other community had some in-person follow-up in Phase 2.

At the phone matching stage, we asked our vendor to send only the "exact" matches that they found (i.e., those that they were most confident about, which meant that we received many fewer "duplicate" matches), which resulted in a lower match rate overall, and many fewer matches for apartments. Address confirmation rate assumptions for Phase 2 were based on Phase 1; although these were much lower than our initial Phase 1 assumptions, they are higher for the listed frames, thus bringing the sample sizes to a reasonable level. Other rates, such as eligibility and completion rates, were based on Phase 1 and were expected to be comparable.

There were also changes to the use of the mail mode. First, as described above, the postcard was hugely unsuccessful and was not used at all in Phase 2. However, although SAQs were abandoned after the first set of replicates in Phase 1, we chose to rely more heavily on the SAQ mailings in Phase 2, bringing this part of the design more closely in line with the original plan. SAQs were mailed to all phone non-matches, as well as to

² We had originally planned to conduct interviews in this Native American community in person, but did not receive permission from the tribal council in time for Phase 1; permission was granted for Phase 2, and we were able to access the local 911 address list to use as our frame.

cases attempted but not contacted and screened, or cases that did not confirm the address, via telephone. A five dollar incentive was also included in each SAQ mailing³.

6. Results of REACH U.S. Survey Design Adjustments

The design changes had an extremely positive impact on data collection efficiency and productivity in Phase 2, and even helped to improve response rates. Phone match rates, while lower than in Phase 1, were more accurate. The listed frames were quite productive, with eligibility rates significantly higher than for the unlisted frames. And the SAQ did remarkably well in Phase 2 and accounted for more than half of its completed interviews. Table 1 below shows some of the significant design changes made between Phase 1 and Phase 2 of REACH U.S.

Table 1: Summary of Design Changes in REACH U.S.

	<i>REACH U.S. Phase 1</i>	<i>REACH U.S. Phase 2</i>
Frame		
Unlisted	Yes (all communities)	Yes (some communities)
Targeted List	Yes (some communities)	Yes (most communities, larger proportion)
RDD	Yes (a few communities)	No
Prior Eligibles	No	Yes (a few communities)
Mode		
Phone	Yes	Yes
Postcard	Yes	No
SAQ	Yes (stopped mid-phase)	Yes
In-person	Yes	No (except for 2 communities)
Address/Phone Match		
Phone Match	Yes (all match types)	Yes (exact matches only)
Address Confirmed	Yes (all communities)	Yes (some communities/household types)

The Phase 1 phone match rates for apartments were much higher than the rates for non-apartments in most communities (Table 2). This was largely because apartments had a greater number of inaccurate phone numbers matched to them. In many cases, we found that apartments in the same building got matched to the same phone number, resulting in duplicates. This inflated the match rate for apartments. When we only considered exact matches—matches that the vendor had confidence in—the match rates for apartments became much lower than the match rates for non-apartments for all communities. Table 3 shows that non-apartment match rates were usually at least double the match rates for apartments.

³ In Phase 1, an experiment was conducted to test various SAQ incentives, and it was decided that a \$5 incentive was the most effective and cost efficient; therefore, this approach was adopted in Phase 2.

Table 2: REACH U.S. Phase 1 Phone Matches by Housing Type

<i>Community</i>	<i>All Matches</i>	
	<i>Apartment</i>	<i>Non-Apartment</i>
A	49%	40%
B	66%	51%
C	51%	58%
D	59%	75%
E	41%	50%

Table 3: REACH U.S. Phase 2 Phone Matches: Exact Matches by Housing Type

<i>Community</i>	<i>All Matches</i>	
	<i>Apartment</i>	<i>Non-Apartment</i>
A	17%	38%
B	22%	45%
C	29%	63%
D	40%	50%
E	22%	41%

Table 4 shows the address confirmation rates for Phase 1 of REACH U.S. As expected, due to fewer exact matches being obtained for apartments, the phone matches for apartments were less accurate, and thus the address confirmation rates for apartments were consistently lower. We did use the community and housing type address confirmation rates, though, to make changes to the use of the address confirmation question in Phase 2. If the rate in a particular community/housing type cell for Phase 1 was 90 percent or higher, we did not ask the question of those households during the screener for Phase 2. This resulted in a shorter screener and, we hoped, one fewer opportunity for respondents to silently refuse participation. For Phase 2, in 16 communities, the address confirmation question was only asked if the housing unit was an apartment (i.e., single-family housing units were not asked the question), and in three additional communities the question was not asked of anyone (all households were exempt). While the address confirmation rates in Phase 2 were lower, due to the fact that we were only asking the question of those more likely *not* to have a correct phone match, we did save screener time for many cases by not asking the question at all.

Table 4: Address Confirmation Rates: Phase 1 of REACH U.S.

<i>Community</i>	<i>Overall</i>	<i>Apartment</i>	<i>Non-Apartment</i>
A	91%	76%	96%
B	86%	73%	89%
C	84%	45%	94%
D	73%	73%	n/a
E	92%	68%	94%

Table 5 shows the increased use of mail questionnaires in Phase 2, by design. It is the mail responses that cover the households that do not have landline telephones, as shown in Table 6 (using Phase 1 data). Thus the Phase 2 design appears to be successful in improving coverage of these difficult-to-reach households based on preliminary Phase 2 data and trends from Phase 1 that we expect to hold in Phase 2.

Table 5: Completion by Mode (Phase 1 vs. Phase 2)

<i>Community</i>	<i>Phase 1</i>		<i>Phase 2</i>	
	<i>Phone</i>	<i>Mail</i>	<i>Phone</i>	<i>Mail</i>
A	76%	24%	24%	76%
B	81%	19%	38%	63%
C	60%	40%	33%	67%
D	80%	20%	33%	67%
E	80%	20%	51%	49%

Table 6: Telephone Status of Phase 1 Households Completing the SAQ

<i>Household Phone Status</i>	<i>% of SAQ Responders</i>
Cell Phone Only	34% (979)
Cell and Landline Phones	44% (1297)
Landline Phone Only	20% (584)
No Phone	2% (63)

7. Conclusions

The REACH surveys have experienced many design and operational changes over the years. Some changes were expected adjustments as experience indicated that design assumptions needed refinement. Where minor tweaks to design changes were not sufficient, we also made major design changes in the middle of the first phase or cycle to achieve the targeted number of completed cases within budget, and the resulting experience informed a revised design for subsequent phases. The largest change, though, was the transition from primarily RDD sampling to address-based sampling to compensate for the reduced coverage of telephone based designs. Address-based sampling required another process of adjusting the design, as initial assumption failed to deliver on the requirements within budget without intervention. Preliminary results from Phase 2 of the address-based sampling approach shows that the revised design, using learnings from the first phase, was more efficient and successful in achieving better coverage of the target population. For future phases of REACH, the design focus will be analysis of potential differences between listed and unlisted telephone households, mode effects, and further improvements in coverage while maintaining efficiency.

References

- Blumberg, S.J. and Luke, J.V. (2010). "Wireless Substitution" Early Release of Estimates From the National Health Interview Survey, July-December 2009." <http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201005.pdf>.
- Harter, R., and Emmons, C. (2003). "The REACH 2010 Community-based surveys and their costs." Proceedings of the American Statistical Association's Joint Statistical Meetings.
- Grady, C., & DeBello, A. (2004). "Conducting telephone survey research in urban, minority communities: An assessment of regional differences in REACH 2010." Presentation at City Futures Conference, Chicago, IL.
- Link, M. W., Battaglia, M. P., Frankel, M.R., Osborn, L., and Mokdad, A. H. (2007). "Reaching the U.S. Cell Phone Generation--Comparison of Cell Phone Survey Results with an Ongoing Landline Telephone Survey." *Public Opinion Quarterly* 2007 71(5):814-839.
- Link, M. W., Battaglia, M. P., Frankel, M.R., Osborn, L., and Mokdad, A. H. (2008). "A Comparison of Address-Based Sampling (ABS) Versus Random-Digit Dialing (RDD) for General Population Surveys." *Public Opinion Quarterly* 2008 72(1):6-27.
- Link, M. W., Daily, G., Shuttles, C. D., Yancey, T. L., and Bourquin, H.C. (2009). "Building A New Foundation: Transitioning To Address Based Sampling After Nearly 30 Years Of RDD." AAPOR 2009 proceedings.
- Murphy J., Burke, A., and Murphy, W. (2003). "Assessing the Validity of Matched Address Data for Unlisted Telephone Numbers in REACH 2010." Proceedings of the American Statistical Association's Joint Statistical Meetings.
- Murphy, W., O'Muircheartaigh, C., Emmons, C.A., Pedlow, S., and Harter, R. (2003). "Optimizing Call Strategies in RDD: Differential Nonresponse Bias and Costs in REACH 2010. Proceedings of the American Statistical Association's Joint Statistical Meetings.
- Tucker, C., Brick, J. M., and Meekins, B. (2007). "Household Telephone Service and Usage Patterns in the United States in 2004: Implications for Telephone Samples." *Public Opinion Quarterly* 2007 71(1):3-22.