Statistical Challenges Facing Cell Phone Surveys

Michael P. Battaglia[1], Martin R. Frankel[1], Ali H. Mokdad[2]
[1] Abt Associates Inc., 55 Wheeler St., Cambridge, MA 02138
[2] University of Washington, Seattle, WA 98105

Abstract
Traditional landline random-digit-dialing (RDD) surveys only include landline telephone numbers, however, the number of households in the U.S. that only have cellular telephone service has grown rapidly in the past few years. One approach to improve the coverage of RDD surveys is to use a dual frame design whereby a sample of cellular telephone numbers is also drawn and a screening instrument is used to identify individuals that only have cellular telephone service. An important issue for this type of sample design is the development of a weighting methodology, especially for state or sub-state surveys.

Key Words: Random-digit-dialing, landline telephone, cell-only households, coverage bias

1. Introduction

In the U.S. it has been well documented (Blumberg and Luke 2008) that while at least 98% of the adult (18+) population currently has telephone access, the proportion of the population that has "cut the cord" in terms of landline telephone service and has adopted cell-phones for voice telephone communication has risen from 4.4% in the first half of 2004 to 14.5% in the second half of 2007. This change in telephone access has particularly impacted younger adult population. The proportion of persons who are cell phone only has risen from 10.3% in the first half of 2004 to 30.6% in the second half of 2007 among persons 18-24. Landline random-digit-dialing (RDD) surveys that only include landline telephone numbers therefore exclude a substantial proportion of the adult U.S. population.

2. Survey Design

The BRFSS collects uniform, state-specific data on preventive health practices and risk behaviors that are linked to morbidity and mortality among adults. The survey is conducted by state health departments with assistance from the Centers for Disease Control and Prevention (further details on the BRFSS survey design, methodology, and questionnaire are available at http://www.cdc.gov/brfss). Three states were selected for participation in the cell phone study: Georgia, New Mexico, and Pennsylvania. The three states were chosen because they represent various geographic regions of the United States and combined they provide a good representation of both the urban/rural and racial/ethnic mix of the U.S. population. The target population for the study consisted of all noninstitutionalized adults aged 18 and older with telephone service living in the three states.

Cell telephone numbers were sampled from dedicated cell phone 1,000 banks and screened for the presence of adults living in private residences within three states. A sample of 23,397 telephone numbers was drawn from cell phone 1,000 banks across the three states. For sample release and management purposes, the sample was divided into random subsamples (replicates) within each state. When a potential respondent was contacted, he or she was first asked questions to determine study eligibility. Specifically, the person answering the telephone was asked whether he or she had been reached on a cell phone, lived in a private residence, was aged 18 years or older, and resided in one of the specified states. For those who responded "No" to any of these questions, the interview was terminated. Those answering "Yes" to all of the screening questions were then asked whether they also had a landline telephone in their home. All of those who had only a cell phone were interviewed, whereas persons who had both types of telephones (dual telephone service) were subsampled for administration of the full interview; the subsampling rate varied depending on state and sample replicate release. The subsampling procedure was used within each state in order to achieve approximately 200 interviews with each of the two cell phone groups. For this study we made an assumption that cell phones are primarily individual devices and not shared household devices; i.e., that there is most often a one-to-one correspondence between the person answering the cell phone (assumed to be

the cell phone owner) and the cell phone itself. As a result, no additional within household selection was undertaken.

A shortened (approximately 12 minutes) version of the 2007 BRFSS core questionnaire was programmed using a computer-assisted telephone interviewing (CATI) system. Interviewing was conducted in English and Spanish. Data were collected from January 18 through April 5, 2007. To increase the response rate and to offset any charges incurred by respondents, a modest payment was provided to respondents. Those who completed the full screening portion of the interview but were not eligible for the interview because of cell or landline subsampling were offered $1. Eligible respondents who completed the detailed interview were offered $10. The questionnaire included questions on interruptions in landline telephone service and cell phone telephone service in the past 12 months. Link et al. (2007) provide further details on the survey methodology.

The cell phone survey was conducted in parallel with the ongoing, monthly BRFSS RDD data collection, thereby facilitating the comparison of results between the two approaches. One adult is randomly selected from a sample household. Telephone survey data from the three participating states for January through March 2007 were used in this analysis. Additional questions were added to the landline telephone survey to determine the type of telephone access in the household (landline and cell phone service or landline service only) facilitating comparison with the cell phone study data. Also, the questionnaire included questions on interruptions in landline telephone service and cell phone telephone service in the past 12 months (not asked in Georgia).

Overall, 1,164 interviews were completed with those reached on a cell phone: 572 interviews with cell phone only households and 592 interviews with households that had both a cell phone and a landline. For the landline survey, 5,788 interviews were completed: 1,972 interviews with persons in landline-only households and 3,816 interviews with those in households with both a landline and a cell phone. As noted earlier for the cell phone survey, all of those having only a cellular telephone were interviewed, while persons having both types of telephones were subsampled for administration of the full interview. The proportion of screened respondents having only cell phones was 35.5 percent in Georgia, 38.5 percent in New Mexico, and 23.6 percent in Pennsylvania. A total of 1,049 potential respondents having both cell phones and landlines were categorized as not selected for interview based on the sub-sampling. To maximize comparability between the landline and cell phone surveys, we used response rate calculations recommended by the American Association for Public Opinion Research. Overall, the cell phone response rate was considerably higher in New Mexico (31.3 percent) than in either Georgia (24.4 percent) or Pennsylvania (23.2 percent). The cell phone response rates were significantly lower than the land-line response rates in New Mexico (51.9 percent versus 31.3 percent) and Pennsylvania (45.3 percent versus 23.2 percent). In Georgia, however, the land-line and cell phone response rates were nearly identical (23.6 percent versus 24.4 percent).

3. Weighting Methodology

The design weight of each completed BRFSS RDD landline interview in a state equals the product of three factors: 1) the reciprocal of the selection probability of the telephone number, 2) the number of adults in the household, and the reciprocal of the number of voice-use landline telephone numbers in the household.

The design weight for the cell phone-only adults interviewed in the cell phone sample in each state equals the reciprocal of the probability of slection of the cell telephone number. The design weight for adults that also have a voice-use landline telephone in their household equals the product of the reciprocal of the probability of slection of the cell telephone number and the reciprocal of the subsampling rate. Within each state a unit nonresponse adjustment was made to the design weights separately for cell phone-only adults and adults with dual phone service.

One can combine the RDD landline interviews with the cell phone-only interviews and poststratify to state-level control totals of age by gender by race/ethnicity. Brick et al. (2007) indicate that one disadvantage of this method is that estimates of the proportion of adults that only have cell phones are likely to be imprecise. Because of differences in telephone usage patterns, adults that only have a cell phone are easier to contact in a cell phone sample than dual service adults, thus the estimated percentage of adults that are cell phone-only will be overestimated. A consequence of this is that estimates of the proportion of adults

who only have cell phone service are considerably higher than the national estimates from the National Health Interview Survey. The AAPOR Cell Phone Task Force points out the problem faced by all state and sub-national surveys: "As a third example, consider a dual frame survey, without overlaps. Here, an RDD frame could be used to interview the landline-only group and those with both types of service, and an RDD cell phone frame could be used for those with only cell service. In this case, weighting adjustments would be made within each frame, and each would be post-stratified so that the sum of the weights for each was proportional to its share of the overall population. However, if the survey is not national in scope, data for such post-stratification may not be available." We developed a statistical methodology to address this issue.

The 2006 National Health Interview Survey provides national and Census Region estimates of the number of households and adults in four telephone usage groups: 1) landline service only (LL), landline and cell telephone service (LLCP), cell phone service only (CP), and no telephone service (NP). A three category dependent variable was created by using the first three categories of telephone usage: Y = 0 (referent group, LLCP), Y = 1 (LL), and Y = 2 (CP). A household-level (HH) multinomial logistic regression was fit using categorical predictor variables (p covariates and a constant term denoted by the vector x) available in the 2006 NHIS, but also available in the 2006 March Current Population Survey (CPS) and the 2006 American Community Survey (ACS): type of living quarters, number of persons in the HH, number of children in the HH, number of elderly adults in the HH, highest education level among the adults in the HH, home tenure status, number of adult males in the HH, number of adult females in the HH, number of adult Hispanics in the household, number of adult nonHispanic blacks in the HH, number of married adults in the HH, and family structure.

A propensity model approach was used. To begin, the coefficients of the NHIS multinomial logistic regression model were applied to the all households in each of the 50 states and the District of Columbia in the CPS except the nontelephone households (the CPS directly identifies nontelephone households). For each of the CPS households three conditional probabilities were calculated: $P(Y=0|x) = P(LLCP)$, $P(Y=1|x) = P(LL)$, and $P(Y=2|x) = P(CP)$. The three conditional probabilities for each household were assigned to all adults in the household.

The BRFSS age by gender by race/ethnicity poststratification weighting is based on state-level control totals provided by Claritas, Inc. For each state we formed Claritas control totals for a 28-category age (7 categories) by gender by race/ethnicity (nonHispanic white versus all other race/ethnic groups) ASR variable. Using the NHIS we estimated the proportion of adults in each of the four telephone usage groups by Census Region. The four estimated proportions for a given Census Region were applied to the Claritas control total for that Census Region to yield telephone usage control totals for each of the four Census Regions that were consistent with the Claritas Census Region controls. We refer to these as the adjusted NHIS control totals.

The next step involved assigned each adult in the CPS to a state by ASR category. Each adult in the CPS has a person weight. In the next step of the propensity modeling, we ratio-adjusted the CPS weights so that the weighted state by ASR distribution of the CPS was in agreement with the Claritas control totals. We refer to this weight as the adjusted CPS weight. Recall that each adult in the CPS either has the three telephone usage conditional probabilities discussed above or lives in a nontelephone household. For the adults in each of the four Census Regions we multiplied the adjusted CPS weight by P(LL). We then ratio-adjusted the adjusted CPS weight so that the weighted number of LL adults in each Census Region was in agreement with the adjusted NHIS LL control totals described above. This weight is referred to as the LL adjusted CPS weight. For the adults in each of the four Census Regions we multiplied the LL adjusted CPS weight by P(LLCP). We then ratio-adjusted the LL adjusted CPS weight so that the weighted number of LLCP adults in each Census Region was in agreement with the adjusted NHIS LLCP control totals described above. For the adults in each of the four Census Regions we multiplied the LLCP adjusted CPS weight by P(CP). We then ratio-adjusted the LLCP adjusted CPS weight so that the weighted number of CP adults in each Census Region was in agreement with the adjusted NHIS CP control totals described above. For the NP adults in each of the four Census Regions we then summed the CP adjusted CPS weight. We then ratio-adjusted the CP adjusted CPS weight so that the weighted number of NP adults in each Census Region was in agreement with the adjusted NHIS NP control totals described above. This iterative process was repeated 30 times in order to develop a final adjusted CPS weight that ensured that the

weighted sample in each state was in close agreement with the Claritas ASR control totals, and for the adults in each Census Region the size of the four telephone usage groups was in close agreement with the adjusted NHIS control totals.

In the last step of the propensity modeling we used the final adjusted CPS weight to estimate the number of adults in each of the four telephone usage groups in each of the three states. Although we used the telephone usage estimates from the CPS in the weighting of the BRFSS landline and cell phone sample, we also applied the propensity modeling to the 2006 ACS as a way of validating the CPS estimates. Table 1 shows the telephone usage estimates, expressed as percentages, for the three states. The CP estimates for the three states are very close.

Given our primary interest in the inclusion of the sample of cell phone-only adults with the landline RDD sample which covers adults that only have landline service and adults that have landline and cell phone service, we limit our discussion to the weighting methodology for combining these two samples.

Recall that the adult respondents in the BRFSS samples in each state have a design weight. The respondents from the two samples were combined. Three final analysis weights were created. The first final weight entailed poststratification to the 28 age by gender by race/ethnicity (ASR) categories using the Claritas control totals. The second weight involved raking the respondents in each state to control totals for two margins: 1) ASR, and 2) the three telephone usage groups (LL, LLCP, and CP) using the estimates developed from the CPS. The third weight was based on raking to three margins: 1) ASR, 2) the three telephone usage groups, and 3) a nontelephone adjustment to account for adults living in households without telephone service. The nontelephone adjustment margin extended the work of Frankel et al. (2003) who used information on interruptions in landline telephone service to compensate for the exclusion of nontelephone households in RDD surveys. BRFSS respondents in each state were classified as interrupted versus not interrupted based on the respondent reporting an interruption in landline telephone service or an interruption in cell phone service in the past 12 months. The raking control totals for each state were constructed from the CPS using (BRFSS proportions were calculated using the design weight):

1. Estimated number of NP adults + estimated number of LL adults x BRFSS proportion of LL respondents with an interruption in landline telephone service) + (estimated number of LLCP adults x BRFSS proportion of LLCP respondents with an interruption in landline telephone service or cell phone service) + (estimated number of CP adults x BRFSS proportion of CP respondents with an interruption in cell phone service).
2. Total estimated number of adults − estimated total from step 1.

4. Findings

The BRFSS questionnaire collected information on health conditions and health risk factors. In Table 2 we compare thirteen prevalence estimates from the RDD landline sample with the combined BRFSS landline and cellular telephone samples. The inclusion of cellular phone-only adults in the prevalence estimates has a varying impact in each of the three states. The absolute value of the mean percent difference is 6.2% in Pennsylvania, 5.7% in Georgia, and 4.1% in New Mexico. Within all three states the inclusion of cellular phone-only adults has an impact on prevalence estimates of health insurance coverage and on the delay of medical treatment due to cost reasons. A large impact on the prevalence estimates is also observed for physical health problems in Georgia, and binge drinking in New Mexico.

The addition of the nontelephone adjustment has a smaller and varying impact on the prevalence estimates. The absolute value of the mean percent difference is 1.0% in New Mexico, 0.3% in both Georgia and Pennsylvania. New Mexico has the largest estimated percentage (4.8%) of adults living in nontelephone households. The largest difference in New Mexico is a 2.4% increase in the prevalence of adults who delayed medical care due to cost reasons.

References

Blumberg, S.J., and Luke, J.V. 2008. Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2007. National Center for Health Statistics.

Brick, J.M., Edwards, W.S., and Lee, S. 2007. Sampling telephone numbers and adults, interview length, and weighting in the California Health Interview Survey cell phone pilot study. *Public Opinion Quarterly*, 71: pp. 793-813.

Frankel, M.R., Srinath, K.P., Hoaglin, D.C., Battaglia, M.P., Smith, P.J., Wright, R.A., and Khare, M. 2003. Adjustments for non-telephone bias in random-digit-dialling surveys. *Statistics in Medicine*, Volume 22, pp. 1611-1626.

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*, 71: pp. 814-839.

Table 1: State Estimates of Telephone Usage for Adults

State and telephone usage group	2006 March CPS	2006 ACS
Georgia:		
Landline-only	27.4%	27.7%
Landline and cell	56.0	55.1
Cell-only	13.5	13.5
Nontelephone	3.1	3.7
New Mexico:		
Landline-only	35.5	36.3
Landline and cell	49.1	48.7
Cell-only	10.6	11.0
Nontelephone	4.8	4.0
Pennsylvania:		
Landline-only	36.8	36.8
Landline and cell	54.7	54.4
Cell-only	7.3	7.5
Nontelephone	1.2	1.4

Table 2: State-level prevalence estimates for various health conditions and behavioral risk factors among adults by single frame versus non-overlapping dual frame design, Behavioral Risk Factor Surveillance System, 2007

State	RDD Landline Sample Prevalence	Combined RDD Landline Sample and Cellular Sample Prevalence	Combined RDD Landline Sample and Cellular Sample Prevalence With Adjustment to Compensate for the Exclusion of Nontelephone Adults	Percent Difference: RDD Landline Versus Combined Samples	Percent Difference: Combined Samples Without Versus With Nontelephone Adjustment
Georgia:					
Current smoking	21.75	24.41	24.61	12.2	0.8
No health insurance	20.76	23.78	23.97	14.5	0.8
Asthma	6.63	6.37	6.38	-3.9	0.2
Binge drinking	22.81	21.87	21.94	-4.1	0.3
Physical health problems	39.50	42.98	43.10	8.8	0.3
Mental health problems	45.37	45.74	45.91	0.8	0.4
Delayed medical treatment due to cost	17.77	19.75	19.86	11.1	0.6
No exercise	27.34	28.63	28.69	4.7	0.2

Diabetes	13.10	13.37	13.34	2.1	-0.2
High blood pressure	32.94	32.96	33.05	0.1	0.3
No influenza vaccination (age 65 years and older)	32.30	34.16	34.11	5.8	-0.1
Limited activity	19.83	20.72	20.79	4.5	0.3
Tested for HIV (age 18-64 years)	51.19	51.60	51.58	0.8	0.0
New Mexico:					
Current smoking	19.87	20.61	20.95	3.7	1.6
No health insurance	21.02	22.82	23.17	8.6	1.5
Asthma	8.76	8.84	8.69	0.9	-1.7
Binge drinking	18.51	20.35	20.59	9.9	1.2
Physical health problems	45.16	44.97	44.69	-0.4	-0.6
Mental health problems	44.71	45.47	45.24	1.7	-0.5
Delayed medical treatment due to cost	16.40	18.14	18.57	10.6	2.4
No exercise	27.33	28.31	28.42	3.6	0.4
Diabetes	9.68	9.98	10.15	3.1	1.7
High blood pressure	26.88	26.52	26.66	-1.3	0.5
No influenza vaccination (age 65 years and older)	28.55	29.87	29.94	4.6	0.2
Limited activity	19.90	19.97	20.19	0.4	1.1
Tested for HIV (age 18-64 years)	34.97	36.70	36.67	4.9	-0.1
Pennsylvania:					

Current smoking	22.28	23.67	23.73	6.2	0.3
No health insurance	9.34	11.26	11.34	20.6	0.7
Asthma	8.71	9.06	9.08	4.0	0.2
Binge drinking	25.42	25.52	25.48	0.4	-0.2
Physical health problems	42.74	43.71	43.86	2.3	0.3
Mental health problems	43.16	46.50	46.67	7.7	0.4
Delayed medical treatment due to cost	9.38	11.05	11.14	17.8	0.8
No exercise	29.40	30.50	30.49	3.7	0.0
Diabetes	8.63	8.96	8.99	3.8	0.3
High blood pressure	27.02	26.81	26.84	-0.8	0.1
No influenza vaccination (age 65 years and older)	23.15	23.70	23.61	2.4	-0.4
Limited activity	18.70	19.61	19.71	4.9	0.5
Tested for HIV (age 18-64 years)	33.47	35.52	35.54	6.1	0.1