

The Impact of a Mixed-mode Data Collection Design on Response and Non-response Bias on a RDD Landline Telephone Survey

S. Sean Hu¹, Carol Pierannunzi¹, Lina Balluz¹

¹Centers for Disease Control and Prevention, 4770 Buford Hwy NE, K-50, Atlanta, GA 30341

Abstract

Nonresponse bias has become a big concern in all telephone surveys conducted in the U.S. Like other RDD telephone surveys, response rates in the Behavioral Risk Factor Surveillance System (BRFSS) have declined steadily since the mid-1990s. In order to improve state-level BRFSS response rates and to assess nonresponse bias, BRFSS is using a mixed-mode data collection design to achieve these goals. A mail follow-up survey to non-respondents of the landline telephone survey was implemented at 4 states in 2010. Characteristics of respondents at the mail follow-up survey (N=2,390) were compared to those from the landline survey (N=11,646). We developed logistic models for each of 14 health indicators to examine whether exclusion of adults from the mail follow-up survey affected estimates after adjusting for demographic characteristics. The extent of the potential for non-response bias in landline telephone surveys was estimated. Using the mail follow-up survey for non-respondents of the landline telephone survey boosted the response rate by 10%. The non-response biases for 3 out of 14 health indicators were slightly higher. Using the mixed-mode design does improve BRFSS response rates.

Key Words: Random-digit-dialing, landline telephone survey, mixed-mode survey design, non-response bias

1. Introduction

The Behavioral Risk Factor Surveillance System (BRFSS) is the world's largest ongoing, list-assisted random-digit-dial (RDD) landline telephone interview health survey, tracking health conditions and risk behaviors in the United States yearly since 1984 (further details on BRFSS are available at <http://www.cdc.gov/brfss>). With support from the Centers for Disease Control and Prevention, the BRFSS is conducted by all 50 state health departments, as well as the District of Columbia, Puerto Rico, Guam, and the US Virgin Islands. Every month, data on behaviors that place health at risk, clinical preventive health practices, and access to and use of health care services are collected from a randomly selected, representative sample of adults aged 18 years and older.

For more than 30 years, Random-Digit-Dialing (RDD) landline telephone survey has provided a cost-efficient strategy for conducting surveys of the US population. During the past decade, however, participation in most RDD telephone surveys has declined, due most likely to factors such as changes in personal communication technologies, growth of call-screening technologies, and heightened privacy concerns in the face of increased telemarketing calls (Steeh, Kirgis, etc., 2001; Curtin, Presser, etc., 2005). Additionally, coverage provided by landline RDD samples has increasingly been called into question. RDD landline frames exclude households that do not have a telephone of any type

(approximately 2.0 percent in 2009) (Blumberg and Luke, 2010). The increased use of cellular telephones has exacerbated this problem with 24.5 percent of households reported to be cell-phone-only (households with no landline telephone) during the second half of 2009 (Link, Battaglia, etc., 2007; Brick, Edwards, etc, 2007; Kuusela, Callegaro, etc, 2008; Blumberg and Luke, 2009, 2010).

As a landline RDD telephone survey, BRFSS had two specific challenges: First, households with cell-phone-only coverage and without residential telephones were not included; therefore, BRFSS might have been excluding persons of lower socioeconomic status (Hu, Balluz, etc., 2011). Second, the survey response rates have been declining over the past several years, making it increasingly difficult to collect survey data using a landline RDD telephone methodology.

Although response rates do not directly measure the degree of nonresponse bias in risk factor and health condition estimates, the demographic and socioeconomic distributions of the design-weighted BRFSS state samples can differ substantially from external sources such as the commercial U.S. demographics data base Claritas, the Current Population Survey, and the American Community Survey. Since some risk factors and health conditions are highly correlated with these demographic and socioeconomic characteristics, there is a potential for nonresponse bias in the estimates. As part of efforts to explore alternative data collection methodologies for the BRFSS, and building on results from the BRFSS Mail Survey Pilot conducted in 2005 and 2006, the Mail Follow-up Survey (MFS) was designed and was implemented in 2010. The MFS was designed to assess the effect of mixed data collection modes—specifically a RDD telephone survey with mail survey follow-up of nonrespondents—on BRFSS response rates, using state-based sampling frames that accurately represent the adult population (Hu, Pierannunzi, and Balluz, 2011). The primary goals of the mail follow-up survey are to increase overall participation in the BRFSS, especially among underrepresented groups including young adult, male, minority, and working populations (Mokdad, Link, etc., 2007; Lambries, Oldendick, etc., 2007) and to assess non-response bias. In addition, results from the MFS will provide information about using mixed data collection modes for conducting cost-effective, reliable surveys of the general public, thus helping to guide future decisions about the conduct of the BRFSS.

2. Methods

2.1 Sample Design

The mail follow-up survey was carried out in selected sample replicates across a set of months. The mail follow-up survey was limited to the landline RDD sample. For the sample replicates selected for the mail follow-up survey the first step was to do a reverse match to obtain residential addresses for a subset of the sample telephone numbers. Based on using commercial address data bases for reverse matching our experience is that 50 to 70 percent of the sample telephone numbers could be matched to a residential address. At this point the sample telephone numbers in the selected replicates can be classified as address-matched versus non-address-matched for three parts of the RDD sample:

- 1) For telephone numbers with an address match, where no adult respondent is selected, the mail questionnaire package was sent to the address resulting from the reverse match after states finish calling attempts to the sample telephone number. The envelope was addressed to [<STATE NAME> Resident] followed by the address. Protocols required

states to rotate the next and last birthday respondent selection techniques to randomly select one adult to fill out the questionnaire.

2) For telephone numbers with an address match, where the adult respondent had been selected and the interview had not been completed (excluding partial interviews), the mail questionnaire package was sent to the address resulting from the reverse match after you finish your calling attempts on the sample telephone number. The envelope was addressed to [<STATE NAME> Resident] followed by the address. The instructions indicated which adult in the household should complete the questionnaire based on the position of the selected adult in the implicit household roster (e.g., oldest male, second oldest female, etc.). For those states that obtained the name of the selected respondent after 15 call back attempts to reach the selected respondent fail to yield an interview, that name may be used in the instructions and in the addressing of the envelope.

3) For telephone numbers without an address match, where the adult respondent was selected and the interview had not been completed (excluding partial interviews), the mail questionnaire package was sent to the address resulting from information obtained from the sample household after states finish calling attempts on the sample telephone number. The envelope was addressed to [<STATE NAME> Resident] followed by the address obtained from the household. The instructions indicated which adult in the household should complete the questionnaire based on the position of the selected adult in the implicit household roster (e.g., oldest male, second oldest female, etc.). For those states that are also willing to obtain the name of the selected respondent after call back attempts fail to yield an interview, that name may be used in the instructions and in the addressing of the envelope.

The mail follow-up survey, just using the core BRFSS questionnaire and without additional modules or questions often added by states, was used as a nonrespondent follow-up technique. It offers the advantage of obtaining the BRFSS interview for all of the telephone respondents in the original sample.

2.2 Data Collection

Initial Questionnaire Mailings: A mail package including a cover letter from state health department, a core BRFSS questionnaire, and a pre-paid business reply envelope was sent by states as they identified potential mail respondents. Mail respondents included those households which have matched addresses and for which the requisite number of telephone calling attempts have been made. In cases where an adult respondent has been selected and/or partial interviews have been completed, the mail package should be mailed out within one week of telephone contact with the household.

Follow-up Contacts: For mail follow-up cases, a series of mail contacts (postcard reminders and second questionnaire mailings) were made. Approximately one week after the first questionnaire mailing, a postcard reminder was sent to all released mail survey cases. The postcard informed households of the importance of the study and the project's toll-free telephone number was provided for households to request another questionnaire mailing (in case the first was misplaced or was never received). In addition, a second questionnaire was mailed four weeks after the original mailing. The second questionnaire was only sent to non-responding cases and it includes a cover letter from the state highlighting the importance of the study. As with the initial questionnaire packages, the follow-up mailings were sent by state health department staff so that the state's postmark appeared on the front.

2.3 Analysis Procedure

We conducted all analyses using SAS, version 9.1 (SAS, Inc.) on unweighted data. The analysis was conducted in three parts. First, we examined the relationship between survey approaches and demographic characteristics of respondents through pairwise contingency tables. Next, a similar approach was used to examine the bivariate relationship between survey approaches and the 15 health conditions and risk behavior measures. Logistic models were developed for each of the health indicators to examine whether the survey approach affected responses after adjusting for the impact of demographic characteristics such as age, gender, education level, marital status, and employment status. Finally, we estimated the potential non-response bias in landline telephone surveys using the metric of relative bias (Cochran, 2007; Lessler and Kalsbeek, 1992). Relative bias assesses by what percentage the survey variables of interest would be overestimated or underestimated if the estimates were based only on the landline survey's respondent data. Values above 0.40 for relative bias changed the 95% level initially established for confidence intervals (Cochran, 2007), which may cause an incorrect inference to be drawn.

3. Results

3.1 Response Rates

Overall, a sample of 152,585 telephone numbers across the 4 states from May to October was drawn. We got a total of 11,646 completed or partially completed landline phone surveys, 6,989 eligible but not interviewed cases, 34,829 unknown eligibility and non-interviewed cases, and 99,121 non-eligible cases from the 4 states (Table 1). The response rate for the landline telephone survey, which we calculated using AAPOR Response Rate Formula 4 (AAPOR, 2010), was 48.2%. After mailing questionnaires to 41,818 non-respondents, we got a total of 2,390 completed or partially completed mail follow-up surveys. The response rate for the combined data of landline phone and mail follow-up surveys was 58.1% using AAPOR Response Rate Formula 4 (AAPOR, 2010).

Table 1: Distribution of Final Disposition Codes in Landline Survey

Category of disposition code	No.
Complete & partial complete	11,646
Non-respondents	
Eligible but not interview	6,989
Unknown eligibility, non-interview	34,829
Not eligible	99,121
Total	152,585

3.2 Demographic Characteristics

The distributions of demographic characteristics among MFS respondents were very close to those of landline telephone survey respondents (Table 2). However, there were still significant differences in demographic distributions between two surveys, including the percentage of respondents who were female (MFS, 54.8%; landline, 60.4%; $p < 0.01$); non-Hispanic white (MFS, 51.7%; landline, 64.3%; $p < 0.01$); 18-34 years old of age (MFS, 6.2%; landline, 10.3%; $p < 0.01$); not working or retired (MFS, 41.1%; landline, 35.8%; $p < 0.01$); married (MFS, 60.8%; landline, 52.7%; $p < 0.01$); high-school or less educated (MFS, 30.6%; landline, 33.8%; $p < 0.01$), and with an annual family income of less than \$35,000 (MFS, 28.7%; landline, 33.1%; $p < 0.01$).

Table 2: Demographics by Survey Mode

Demographic Characteristic		Landline (N = 11646)	Mail Follow-up (N = 2390)	χ^2 p value
Gender	Male	39.6	41.8	<0.01
	Female	60.4	54.8	
	Unknown	0.0	3.4	
Ethnicity	White	64.3	51.7	<0.01
	Black	4.8	3.6	
	Hispanic	5.5	3.3	
	Asian	13.8	36.6	
	Other races	10.3	3.2	
	Unknown	1.2	1.6	
Age	Age 18 to 24	3.0	1.8	<0.01
	Age 25 to 34	7.3	4.4	
	Age 35 to 44	13.5	9.7	
	Age 45 to 54	18.9	18.6	
	Age 55 to 64	23.5	24.2	
	Age 65 or older	32.4	38.9	
	Unknown	1.4	2.4	
Education	< High School	7.1	6.7	<0.01
	High School	26.7	23.9	
	Some College	26.1	28.5	
	College or more	39.7	39.7	
	Unknown	0.4	1.3	
Income	<\$10,000	4.3	4.6	<0.01
	\$10,000 - \$14,999	4.8	5.2	
	\$15,000 - \$19,999	6.3	5.1	
	\$20,000 - \$24,999	8.0	5.5	
	\$25,000 - \$34,999	9.7	8.3	
	\$35,000 - \$49,999	13.0	12.9	
	\$50,000 - \$74,999	14.0	15.5	
	\$75,000+	26.9	30.5	
Unknown	6.2	12.5		
Employment	Employed for wages	42.9	40.5	<0.01
	Self-employed	7.9	8.4	
	Not work	6.8	3.7	
	A homemaker	11.2	6.9	
	A student	1.7	1.3	
	Retire	29.0	37.4	
	Unknown	0.4	1.8	
Marital Status	Married	52.7	60.8	<0.01
	Divorced	13.9	10.8	
	Widowed	13.4	12.6	
	Separated	1.9	0.9	
	Never married	15.3	12.1	
	Member of unmarried couple	2.2	2.0	
	Unknown	0.6	0.8	

3.3 Key Health Conditions and Risk Factor Indices

We found that the landline survey produced significantly higher prevalence estimates than the MFS for ever having asthma, current smoking, ever being tested for HIV, having teeth cleaned, and ever having a sigmoidoscopy exam (Table 3). In contrast, the respondents in the MFS reported significantly higher prevalence estimates than those in the landline for not receiving care due to cost, having a flu shot, and ever having a mammogram. These findings are not surprising, given the difference in demographic distribution between the two survey approaches.

However, after we used logistic regression to adjust for other potential confounders, including age, gender, ethnicity, education level, marital status, and employment status, only three of these differences persisted. When comparing the respondents in the MFS with those in the landline survey, the odds of a “yes” response was lower by 23% for having teeth cleaned in the past 12 months. In contrast, the odds of a “yes” response was greater by 42.3% for not receiving care due to cost, by 20.4% for ever being tested for HIV.

For 2 of 3 health risk and health condition indices that remained different between surveys after controlling demographics, the relative biases were negative values, indicating that the landline survey underestimated prevalence by 2.0% for not receiving care due to cost and by 8.5% for having a flu shot in the past 12 months. In contrast, one relative bias estimate was a positive value, indicating that the landline survey overestimated prevalence by 1.0% for having teeth cleaned in the past 12 months.

Table 3: Prevalence Estimates for Various Health Conditions and Risk Factors, by Survey Methods, and Adjusted Odds Ratios for Comparison of Survey Methods

Health Conditions/Risk Factors	Prevalence Estimate ^a			Adjusted Odds Ratio ^b			Relative Bias
	Landline (N=11,646)	Mail Follow-up (N=2,390)	χ^2 p value	Mail Follow-up vs. Landline			
				OR	95% of CI		
Feeling Good or Better Health Status	82.8	83.2	>0.05	1.012	0.885	1.158	
Any kind of health care coverage	94.4	94.8	>0.05	0.822	0.662	1.020	
Not received care due to cost	7.8	8.1	<0.05	1.423**	1.192	1.698	-1.996
Ever had angina or coronary heart disease	5.8	6.8	>0.05	1.213	0.995	1.478	
Ever had asthma	15.6	13.2	<0.01	0.952	0.829	1.094	
Current cigarette smokers	14.9	11.7	<0.01	0.937	0.798	1.100	
Doing physical activities or exercise	78.8	78.2	>0.05	0.945	0.841	1.063	
Ever tested for HIV	37.3	22.5	<0.01	0.983	0.839	1.153	
Having teeth cleaned in the past 12 months	76.3	74.9	<0.05	0.770**	0.687	0.864	0.952
Had a flu shot in the past 12 months	52.2	60.8	<0.01	1.204**	1.090	1.330	-8.548
Ever had a mammogram	83.3	86.3	<0.01	0.971	0.772	1.221	
Ever had a Pap test	94.4	95.1	>0.05	1.308	0.958	1.784	
Ever had a PSA test	69.6	66.3	>0.05	0.990	0.811	1.033	
Ever had a Sigmoidoscopy exam	69.6	62.5	<0.01	0.914	0.808	1.033	

Note: Data was unweighted.

* Significant at $p < 0.05$; **Significant at $p < 0.01$

4. Conclusion

Our results showed that using mail follow-up survey to non-respondents of the landline survey did increase response rate by 10%. There was no substantial difference among demographic characteristics between landline and mail follow-up respondents. However, the mail follow-up survey failed to reach the under-represented groups of landline survey

respondents such as male, young adult, and minority population. In addition, mail follow-up respondents only differ from landline respondents on a few of risk behaviors and health conditions. Non-response bias is small and could be addressed by appropriate weighting.

Using the mixed-mode design does lead to higher response, but compared with a single-mode design, there are several problems. The use of multiple modes may raise issues of comparability across modes. For example, questions asked by an interviewer over the telephone, as opposed to being asked on paper, may be more likely to invoke socially desirable responses (Turner, Ku, etc, 1998). Evidence exists that survey mode can affect respondents' answers to questions, even when questions are identically worded (Dillman, Christian, etc., 2005). Furthermore, questions asked on paper are more likely to ensure privacy and allow the respondent to complete the survey at his or her convenience. However, complex forms in which some questions are to be skipped cannot be used, and literacy issues must be considered regarding the MFS.

The study has two limitations. Since the mail follow-up survey was only piloted in 2-4 months in 4 states, the sample size was small. Furthermore, the data from the mail follow-up survey was not linked back to the landline survey's data when we conducted analysis. So, the combined data of landline and mail follow-up surveys was unweighted for data analyses.

Acknowledgements

The authors thank the Behavioral Risk Factor Surveillance System state coordinators at 4 states for their critical assistance in the design and execution of this study. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

References

19. American Association for Public Opinion Research. 2009. Standard Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys. Lenexa, KS: American Association for Public Opinion Research; 2009 (http://www.aapor.org/AM/Template.cfm?Section=Standard_Definitions&Template=/CM/ContentDisplay.cfm&ContentID=1819)(Accessed April 26, 2011).
- Blumberg SJ, Luke JV. 2009. Reevaluating the need for concern regarding noncoverage bias in landline surveys. *Am J Public Health*, 99(10):1806-10. PMID: 19696381.
- Blumberg SJ, Luke, JV. 2010. Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July-December 2009. National Center for Health Statistics. Available from: <http://www.cdc.gov/nchs/nhis.htm> [Accessed March 8, 2011].
- Brick JM, Edwards WS, 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(5):793-813. doi:10.1093/poq/nfm052.
- Cochran WG. Sampling Techniques. 3rd ed. 1977. New York, NY: Wiley.

- Curtin R, Presser S, and Singer E. 2005. Changes in Telephone Survey Nonresponse over the Past Quarter Century. *Public Opinion Quarterly*, 69(1):87-98.
- Dillman DA, Christian LM. 2005. Survey modes as a source of instability in responses across surveys. *Field Methods*, 17(1):30-52.
- Hu SS, Balluz L, Battaglia MP, Frankel MR. 2010. The impact of cell phones on public health surveillance. *Bulletin World Health Organization* 2010;88(11):799.
- Hu SS, Balluz L, Battaglia MP, Frankel MR. 2011. Improving public health surveillance using a dual frame survey of landline and cell phone numbers. *Am J Epidemiology*, 173(6):703-11.
- Hu, SS, Pierannunzi, C., Balluz, L 2011. Integrating a Multi-mode Design Into a National Random-Digit-Dialing Telephone Survey. *Preventing Chronic Disease*, 9.
- Kuusela V, Callegaro M, Vehovar V. 2008. The influence of mobile telephones on telephone surveys, Chapter 4 in Lepkowski, J.M., Tucker, C., Brick, J.M., de Leeuw, E.D., Japac, L., Lavrakas, P. J., Link, M. W., Sangster, R. L. (Eds.) *Advances in Telephone Survey Methodology*. New York, NY: Wiley; 87-112.
- Lambries D, Oldendick R, Link MW, Williams L. 2007. Young people may not be in jail, but most are in cells: cross-time impact of changes in phone use on telephone survey representativeness. *Public Opinion Quarterly*, 71(3):E485-E518; doi:10.1093/poq/nfm013.
- Lessler JT, Kalsbeek WD. *Nonsampling 1992. Error in Surveys*. New York, NY: Wiley.
- Link M, Battaglia M, Frankel M, Osborn L, Mokdad A. 2007. Reaching the US cell phone generation: comparison of cell phone survey results with an ongoing landline telephone survey. *Public Opinion Quarterly*, 71(5):814-839. doi:10.1093/poq/nfm051.
- Mokdad AH, Link MW, Hu SS. 2007. Where are we losing respondents? Trends in survey eligibility and participation. *Public Opinion Quarterly*, 71(3):E485-E518; doi:10.1093/poq/nfm013.
- Steeh, Charlotte, Nicole Kirgis, Brian Cannon, and Jeff DeWitt. 2001. "Are They Really as Bad as They Seem? Nonresponse Rates at the End of the Twentieth Century." *Journal of Official Statistics* 17, 2: 227-247.
- Turner, C., L Ku, and S. Rogers. 1998. "Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology." *Science*, 280:867-73.