

## Tradeoffs in Quality: Examining the Relationship between Cell Telephone Respondent Location and Data Quality

Christopher Ward<sup>1</sup>, Becky Reimer<sup>1</sup>, Laurie Elam-Evans<sup>2</sup>, David Yankey<sup>2</sup>, Meena Khare<sup>3</sup>

<sup>1</sup>NORC at the University of Chicago

<sup>2</sup>National Center for Immunization and Respiratory Diseases, CDC

<sup>3</sup>National Center for Health Statistics, CDC

### Abstract\*

A dual-frame telephone survey including a cell telephone frame poses multiple challenges to survey researchers. It is becoming increasingly important to include cell telephone samples as the proportion of cell telephone-only and cell telephone-mainly households continues to increase in the United States. However, cell telephone samples often have lower response rates and higher operational costs than landline telephone samples. This research examined another dimension affected by the inclusion of cell-phone frames: data quality. In particular, we examined whether data quality varied by respondents' sample frame and location (landline telephone-at-home, cell telephone-at-home, and cell telephone-away-from-home) at the time of the interview. We used data from the 2012 National Immunization Survey (NIS), a national, dual-frame random digit dial survey (analytic subsample:  $n = 7,570$ ). Three types of data were evaluated to examine the relationship between sample frame, respondent location, and data quality: the completeness of questionnaire items (e.g., reporting on vaccinations their children had received), completeness of reporting on socio-demographic characteristics, and ability to report healthcare provider names and contact information. Results from analyses showed that data quality on some variables may have been lower for respondents who were away from home at the time of the interview than those who were at home. This research provides insight into the differential reporting of certain information provided by cell telephone-away respondents and the conditions under which these respondents may be most likely to provide complete, accurate data.

**Keywords:** data quality, cooperation, cell telephone, landline, CATI, location, interview, dual-frame, home, away

### Background

A dual-frame telephone survey including a cell telephone frame poses multiple challenges to survey researchers. Cell telephone samples have become integral to computer-assisted telephone interviewing (CATI) over the past several years (Curtin, Presser, & Singer, 2005), because the proportion of cell-only and cell-mainly households has grown at about 2-3% per year over past decade (Blumberg & Luke, 2013). In response to this increasing proportion of cell users in the population, the inclusion of a cell telephone frame in national CATI surveys generally increases coverage of cell-only users and is expected to reduce potential noncoverage bias in estimates (Keeter et al., 2007). The benefits of cell telephone interviewing are not, however, without tradeoffs. In comparison to landline respondents, cell telephone respondents are often more difficult to contact, pose challenges for interviewer engagement, and tend to be more expensive to interview (Link et al., 2007). These operational challenges often pose greater costs to design surveys with cell telephone

\*The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the National Center for Health Statistics, Centers for Disease Control and Prevention, or NORC at the University of Chicago.

samples, as they tend to require more call attempts to complete an interview, additional costs for refusal aversion training, and additional lines of sample to compensate for low response rates (CDC, 2013).

There are potential additional burdens to cell telephone respondents who are away from home at the time of interviewing. For example, privacy concerns are heightened for some, as the interview may take place in public. Likewise, cognitive burden may be greater due to the distractions of multitasking. Previous research suggests that at least part of the increased challenge of interviewing cell telephone respondents lies in attempting to contact and interview them while they are away from home (Ward et al., 2013).

Further compounding these challenges is the possibility that cell telephone respondent's location may affect the data quality, which has been addressed briefly in the literature (Lavrakas, Tompson, & Benford, 2010). This study investigates the extent to which the location of cell respondents at the time of interview may be affecting the data quality. Data quality was investigated across four domains: the extent of missingness in respondent-reported behaviors and sociodemographic characteristics, the extent and quality of respondent-reported healthcare provider contact information, operational performance rates, and child vaccination rates.

### **Research Data and Methods**

Analyses were performed on data from the 2012 National Immunization Survey, a large, national- and state-level, dual-frame RDD telephone survey sponsored by the Centers for Disease Control and Prevention (CDC, 2013). To collect data about childhood vaccinations, the NIS screens for households with at least one child aged 19 - 35 months and uses a take-all household approach (i.e., the study does not screen to determine the extent to which a respondent owned or used one or more cell or landline telephones). The 2012 NIS included a question that is central to our study: "Would you mind telling me if I reached you today away from home or at home?" The structure of the survey instrument is outlined in Table 1.

The location of cell-phone respondents (whether at home—"cell-home"—or away from home—"cell-away") at the time of interview was collected in the demographic section of the survey (Section C). Landline respondents (including users of Internet telephones or home Wi-Fi) were presumed to be at home at the time of interview and therefore not asked the question. A range of sociodemographic questions used in this analysis were also collected in Section C, such as the race and ethnicity of the child, the mother's education, and household income.

Section D of the NIS contains the Provider Look-up (PLU), which enables the interviewer to collect contact information for the child's vaccination provider(s). Interviewers are trained to use the PLU tool to probe respondents for contact information on all vaccination providers for the child or children in the household. As the interviewer interacts with the PLU, he or she may modify contact information fields from retrieved providers or enter entirely new providers as necessary. Interviewers may find a respondent's vaccination provider through the course of conversation, find the provider but modify certain contact information fields, or enter all of a provider's contact information at once.

Obtaining vaccination provider contact information is crucial to the success of the NIS, as it is the mechanism by which the study obtains vaccination histories for children. These providers are mailed an Immunization History Questionnaire (IHQs) to collect vaccination histories. The data collection process for IHQs, called the Provider Record Check (PRC), is outlined in Figure 1.

The NIS-PRC is a complex data collection effort. The efficient movement of IHQs through the NIS-PRC helps to ensure high data quality and accurate representation of children from the CATI portion of the NIS. Each of the processes depicted in Figure 1 represents a point at which inefficiency may be introduced. For example, when a respondent fails to give sufficient contact information for a child's medical provider to be mailed an IHQ (hereafter, to be "mailable"), the provider's correct contact information must be verified or obtained through a locating process (e.g. through using health provider databases, yellow pages, or the Internet). This locating process, while necessary, adds an additional layer of inefficiency. If a provider returns an IHQ but reports that the child was not in his or her care, no vaccination information can be obtained. The analysis of these data collection processes attempts to identify whether respondent location is associated with data quality.

### **Analytic Subsample**

The analytic subsample included only respondents who completed the screener or household interview questionnaires, and granted consent to contact vaccination providers on one call attempt. This criterion also ensured that respondent behavior during the interview (that is, from screening to completion) could be linked with his or her location. To ensure comparability across households, the subsample is also limited to the first child named in multiple-child households. In addition, to further simplify the analysis of cases through NIS-PRC, the subsample was limited only to households reporting a single vaccination provider. Twenty-six interviews where cell-phone respondents reported "don't know" or "refused" to the location question at the time of interview were excluded from analyses. The analytic subsample included the 7,570 cases (approximately 30 percent of the total 2012 sample of 25,334) that met the inclusion criteria. The distribution of respondent location status as well as the unweighted distribution of demographic characteristics of the subsample are presented in Tables 2 and 3.

### **Analysis**

To investigate the relationship between data quality and location of the cell telephone respondent, rates of missingness (responses of "don't know" or "refused") were compared on key socio-demographic and behavioral variables. Rates of missingness were also compared for each of the vaccination provider address and name fields. With landline as a reference, pairwise tests of proportions were computed on each variable to determine whether missingness varied significantly by respondent location.

To understand the operational consequences arising from data quality concerns, a second set of comparisons was conducted on operational performance rates and vaccination rates. With landline as a reference, pairwise tests of proportions were computed on each performance measure by respondent location to determine the consequences respondent location may have for operational efficiency and selected vaccination rates.

All analyses accounted for the complex sample design of the NIS, and all results reflect the use of the NIS base weight (the inverse of the probability of sample selection). Unlike the final NIS weights, the base weights do not incorporate any noncoverage, nonresponse, or other adjustments. The base weight was selected because the nonresponse and noncoverage adjustments on other NIS weights could spuriously interfere with observing item nonresponse, a key outcome in this study.

## **Results**

### **Rates of Missingness**

Table 4 compares rates of missingness across a range of sociodemographic variables. Rates of missingness were overall very low across the majority of variables examined. The rate of missingness varied significantly on three of eight examined variables: child's ethnicity, mother's education, and the exact household income question (that is, the first income question asked before a series of bracketed income questions). The most notable difference was that on exact household income, where the rate of missingness for cell-away respondents was more than double that of the landline and cell-home subsamples. In contrast, the rate of missingness on receipt of benefits from the Women, Infants, and Children (WIC) program was significantly lower for cell-away respondents than for landline respondents. In general, data quality was reasonably consistent across these items with the notable exception of the rate of reporting exact household income.

### **Performance Rates in NIS-Provider Record Check**

Table 5 compares rates of missingness on provider's name, clinic name, and address fields, all of which were furnished by respondents. A majority (54.75%) of cell-away respondents did not give provider name information in the interview, a rate that was significantly higher than the landline respondents who were unable to do so. About 5% of cell-away respondents were unable to report clinic name, which was significantly lower than the landline respondents (10.34%) without such information. Cell-home respondents were significantly more likely to omit address information than were landline respondents, and cell-away respondents were significantly more likely to omit zip code information. Differences across location may in part be attributable to the availability of health care and vaccination records at home, whereas cell-away respondents do not ordinarily have access to these documents.

Table 6 compares key performance rates in the NIS Provider Record Check by respondent location status. These rates are important indicators of overall data collection execution and efficiency. The rate at which cell-away respondents report sufficient details about healthcare provider contact information to have a mailable address is significantly lower – by almost thirty percentage points – than for landline respondents. This is a serious operational inefficiency, as all providers without mailable contact information must be subsequently located and identified by data collection clerks to find their proper address. However, no significant differences were observed on IHQ return rate or on the rate at which providers reported no record of having cared for the child. The conditional adequate provider data rate – that is, the proportion of children for whom sufficient vaccination history was collected – is significantly lower among cell-home respondents than among landline respondents, but no significant difference was observed for cell-away.

Despite the significantly differential operational performance rates on these measures, Table 7 shows no significant pairwise differences were observed on either type of vaccination rate. That is, while differences were observed on the operational side, rates on a key variable of interest (vaccination) did not vary by respondent location.

### **Limitations**

By the nature of its structure, the restricted analytical subsample may not be generalizable to all respondents in the 2012 NIS or to other telephone surveys. Only respondents who were screened, completed the interview, identified one vaccination provider, and granted consent to contact that provider on one call attempt were included in the sample. This criterion may bias the results toward the behavior of the most cooperative respondents, as it excludes respondents who ever broke off during the interview in the course of data collection. As a consequence, variation in missingness is likely greater among all respondents.

Further compounding this effect is the exclusion of respondents who did not know or refused to respond to the cell location question. Because item nonresponse is often correlated across various questions, the exclusion of these respondents may have reduced variation in “don’t know” and “refused” responses in this analysis. Therefore, these results may underestimate the relationship between respondent location and data quality.

Finally, rates of missingness on some items may interact in unexpected ways with the item topic. Consider the item that asks the respondent whether his or her child ever received benefits from the WIC program, which aims to assist nutritionally at-risk children as well as children from low-income families. On the one hand, reporting receipt of WIC may introduce privacy concerns, especially in public spaces while away from home. On the other hand, unexpectedly, the rate of missingness for cell-away respondents was the lowest among the three location groups. Additional research may be needed to explain this finding.

### **Conclusions and Future Research**

This analysis provides some evidence that the quality of data provided by a respondent may vary with his or her location during the interview. Significantly different responses to the income question suggest that question sensitivity or measurement error may vary by location. These results are consistent with previous research on the behavior of cell telephone users away from home.

Obtaining vaccination histories from providers is central to the purpose of the NIS, but it appears that cell respondents away from home may have some additional difficulties in reporting provider contact information. Additionally, providers that do not have readily available contact information must undergo an additional locating process, thus introducing operational inefficiency. Although cell-away respondents’ difficulty in recalling provider contact information may somewhat hinder the operations of the PRC data collection effort, their responses do not appear to have a significant impact on up-to-date rates or the rate at which providers return the IHQ.

Further research should examine the unique cognitive burdens placed on cell respondents away from home and the behavioral consequences. Previous research has speculated that privacy concerns – particularly on very sensitive topics – and increased cognitive burden

may have an impact on these respondents, and this research largely corroborates that work. In some measures of data quality, being away from home was associated with lower-quality data (such as missing exact household income information), whereas observed differences in other measures (such as reporting provider address) arose between landline and cell (both away and at home) respondents. Future research and investigations into questionnaire design should examine how to mitigate the challenges of interviewing cell telephone respondents, especially those away from home during the interview. Because the inclusion of both landline and cell telephone frames is so vital to sustaining high coverage rates, it is necessary to understand better how to obtain high-quality data from cell-away respondents in light of the higher costs of cell telephone interviewing.

### References

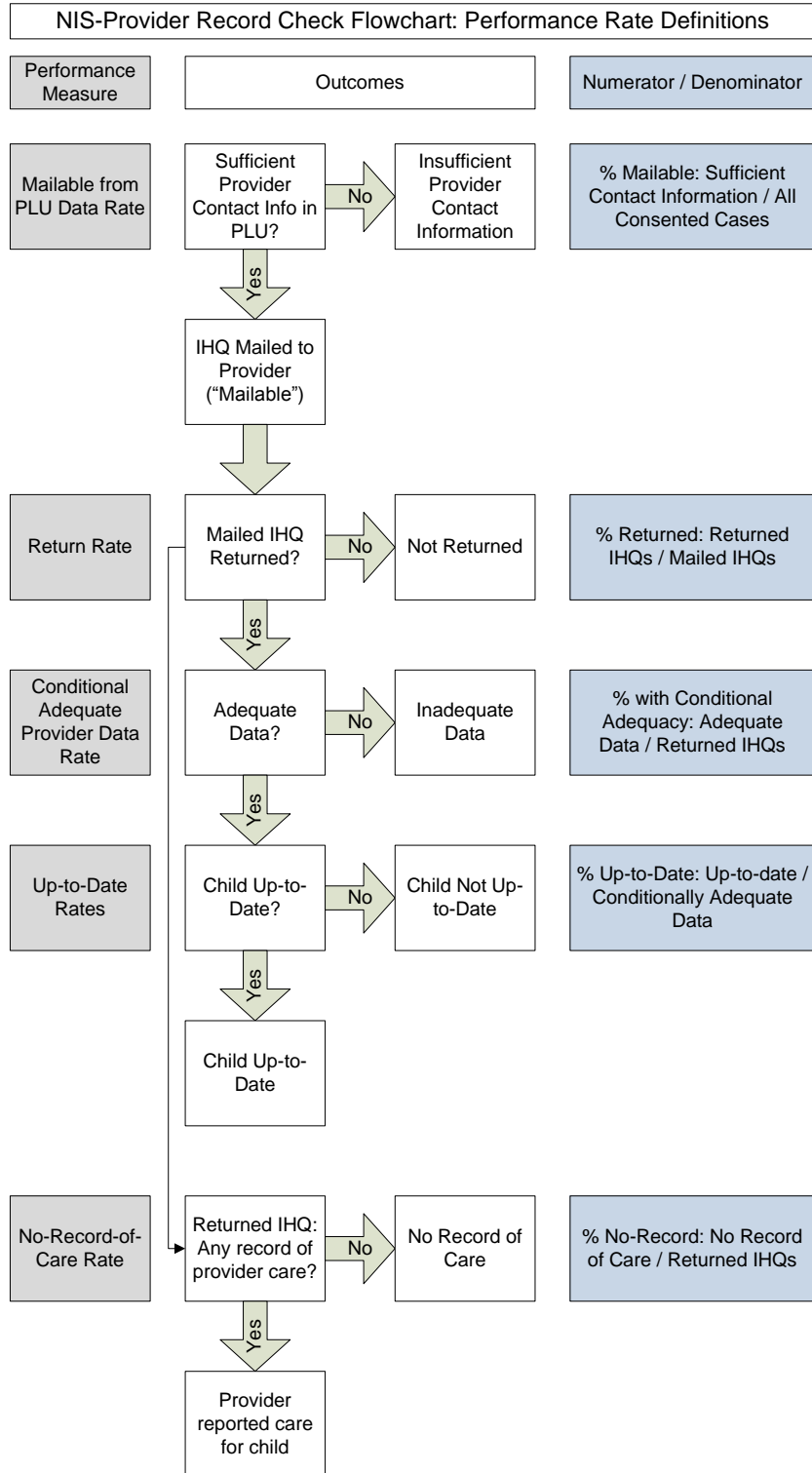
- Blumberg, S.J., Luke, J.V. (2013). *Wireless substitution: early release of estimates from the National Health Interview Survey, January–June 2013*. Hyattsville, MD: National Center for Health Statistics. Retrieved May 12, 2014, from <http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201312.pdf>.
- Centers for Disease Control and Prevention (2013). *The 2012 National Immunization Survey Data User's Guide*. Atlanta, GA: Centers for Disease Control and Prevention. Available at: [http://www.cdc.gov/nchs/nis/data\\_files.htm](http://www.cdc.gov/nchs/nis/data_files.htm). Retrieved June 25, 2014.
- Curtin, R., Presser, S., Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly*, 69:87–98.
- Keeter, S., Kennedy, C., Clark, A., Tompson, T., Mokrzycki, M. (2007). What's Missing from National Landline RDD Surveys? The Impact of the Growing Cell-Only Population. *Public Opinion Quarterly*, 71:772–792.
- Lavrakas, P.J., Tompson, T.N., Benford, R. (2010). Investigating Data Quality in Cell Phone Surveying. Paper presented at the 65th annual conference of the American Association for Public Opinion Research; Chicago, IL.
- Link, M.W., Battaglia, M.P., Frankel, M.R., Osborn, L., 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*, 69:814–39.
- Ward, C.D., Reimer, B., Khare, M., Black, C. (2013). Home Is Where the Cooperation Is: The Association between Interview Location and Cooperation among Cell-Phone Users. In *JSM Proceedings*, Survey Research Methods Section. Alexandria, VA: American Statistical Association.

## Appendix

**Table 1: Structure of the NIS CATI Instrument, 2012**

<b>Section Title</b>	<b>Topic</b>
Section S	Screeners for age-eligible child(ren)
Section B	Child(ren)'s Vaccination History from the respondent
Section C	Demographic and Socioeconomic Information  This section includes the question about whether cell respondent is at or away from home at the time of interview. A completed Section C defines a completed interview.
Section D	Healthcare Provider Look-up Tool: Precursor to the NIS Provider Record Check  This section includes the question asking the respondent to grant consent to contact the child's health care provider(s) to obtain vaccination records.
HIM	Health Insurance Module

**Figure 1: NIS Provider Record Check, 2012**





**Table 2: Respondent Location**

Location Type	Landline	Cell-Home	Cell-Away
<b>Interviewed with Consent Subsample (n)</b>	4,005	2,600	965
<b>Proportion of Subsample (%)</b>	52.91	34.35	12.75

**Table 3: Selected Respondent Demographics by Location**

Demographic Variable	Landline	Cell-Home	Cell-Away
<b>Respondent Is Mother (%)</b>	74.53	63.11	75.88
<b>Mother's Race/Ethnicity: Hispanic (%)</b>	8.64	11.27	14.20
<b>Mother's Race/Ethnicity: Non-Hispanic White (%)</b>	74.76	66.32	71.38
<b>Mother's Race/Ethnicity: Black (%)</b>	11.25	11.61	10.38
<b>Mother's Race/Ethnicity: Other (%)</b>	6.49	7.88	6.96
<b>Mother's Age: ≤ 29 (%)</b>	26.84	45.70	44.19
<b>Mother's Age: &gt; 29 (%)</b>	73.16	54.30	55.81

**Table 4: Questionnaire Items: Rates of Missingness (%)**

Items	Landline n=4,005	Cell-Home n=2,600	Cell-Away n=965	Overall n=7,570
<b>Child's Ethnicity (Hispanic/Latino)</b>	0.19	0.04	0.28	0.15
<b>Child's Race (White/Black/Other)</b>	0.4	0.65	0.09	0.44
<b>Mother's Education</b>	0.48	0.18	0.70	0.40
<b>Mother's Marital Status</b>	0.40	0.06	0.19	0.19
<b>Exact Household Reported Income</b>	21.4	17.66	44.21*	25.59
<b>Reported Ever Received WIC</b>	0.17	0.62	0.02*	0.34
<b>Child Ever Breastfed</b>	0.17	0.14	0.13	0.15
<b>Child Received Influenza Vaccination</b>	1.61	1.25	1.07	1.30

\*:  $p < 0.05$ ; \*\*:  $p < 0.01$

**Table 5: Provider Fields: Rates of Missingness (%)**

Performance Item	Landline	Cell-Home	Cell-Away	Overall
Provider Name	27.46	31.67	54.75*	37.14
Clinic Name	10.34	10.18	4.96**	8.73
Address	4.35	6.48*	9.86	6.87
City	0.70	0.42	0.23	0.44
State	1.83	2.66	7.13	3.71
Zip Code	5.94	5.65	35.25*	14.22

\*:  $p < 0.05$ ; \*\*:  $p < 0.01$ **Table 6: PRC Performance Rates: Households with Consent (%)**

Performance Measure	Landline	Cell-Home	Cell-Away	Overall
Rate: IHQs Mailable before Any Staff Locating Effort	71.24	71.65	43.48*	64.22
IHQ Return Rate	94.19	92.18	89.45	92.05
Conditional Adequate Provider Data Rate	95.05	92.45*	89.52	92.44
Rate: Returned IHQ: No Record of Care	3.17	3.33	1.73	2.88

\*:  $p < 0.05$ ; \*\*:  $p < 0.01$ **Table 7: Up-to-Date Vaccination Rates by Respondent Location (%)<sup>†</sup>**

Comparison of Vaccination Rates	Landline	Cell-Home	Cell-Away	Overall
Up-to-Date on Vaccine Series 4:3:1:3:3:1 <sup>‡</sup>	80.13	77.16	86.09	80.22
Up-to-Date on Vaccine Series 4:3:1:3:3:1:4 <sup>§</sup>	74.87	73.64	84.17	76.59

<sup>†</sup>Results are base-weighted and do not reflect official NIS vaccination estimates<sup>‡</sup>4:3:1:3:3:1 series includes the following provider-reported vaccines:

4 or more doses of diphtheria, tetanus, pertussis vaccine (DTP/DT/DTaP)

3 or more doses of poliovirus vaccine

1 or more doses of measles antigen-containing vaccine (MCV)

3 or more doses of haemophilus influenzae type b (Hib) vaccine

3 or more doses of hepatitis B vaccine

1 or more doses of varicella vaccine

<sup>§</sup>4:3:1:3:3:1:4 series includes 4:3:1:3:3:1 series plus 4 doses of pneumococcal conjugate vaccine (PCV)