Published In *JSM Proceedings,* Survey Research Methods Section. Alexandria, VA: American Statistical Association.

Using Lists with RDD Samples: An examination of bias, cost and variance estimates

Anthony M. Roman[1], Susan M. Rogers[2], Charles F. Turner[3], Elizabeth Eggleston[2], James R. Chromy[4], Sylvia Tan[2]

[1] Center for Survey Research, University of Massachusetts-Boston, 100 Morrissey Blvd., Boston, MA 02125
[2] Program in Health and Behavioral Measurement, Research Triangle Institute, 701 13th St. NW, Suite 750, Washington, D.C. 20005
[3] City University of New York, Queens College and Graduate Center, Flushing, NY 11367
[4] Statistics and Epidemiology Division, Research Triangle Institute, Research Triangle Park, NC 27709

Background

The Monitoring STIs Survey Program (MSSP) was a telephone survey conducted in Baltimore, Maryland from 2006 through 2009. It was a coordinated effort among the Research Triangle Institute (RTI), the Center for Survey Research of the University of Massachusetts-Boston (CSR), the University of North Carolina-Chapel Hill (UNC) and Johns Hopkins University (JHU). The survey screened households in Baltimore for people 15-35 years of age and then randomly selected one age-eligible person for interview. The interview itself was conducted by computer using Telephone Audio Computer-Assisted Self Interviewing (TACASI). Respondents were paid $20 for completing the interview and were asked to participate further by submitting a biospecimen that would be tested for three specific sexually transmitted infections (STIs), namely gonorrhea, chlamydial infection, and trichomoniasis. For those who agreed, a specimen collection kit was mailed to the respondent along with instructions for use of the kit and for mailing the specimen to the lab for testing. All postage was prepaid so the respondent incurred no costs for performing this process. If the respondent supplied a specimen to the lab, they were sent an additional $40 to $100. If the lab results were positive, respondents were notified of their result and offered further examination and free treatment at one of the city's public STI clinics. In addition, the Baltimore City Health Department (BCHD) was notified of positive results for gonorrhea and chlamydial infection as required by Maryland law. For selected respondents under 18 years of age, parental permission was first obtained over the telephone. Parental permission was requested separately for the interview and submission of the biospecimen. Once permission was obtained, attempts were made to contact the minor and complete the interview. Only the minor was notified of their positive test result, not the parent.

Sample Design and Data Collection Protocol

The original sample design called for a pure random-digit-dialed (RDD) sample of all telephone numbers serving the city of Baltimore, Maryland. For a survey to be conducted over three years from 2006 to 2009, the pure RDD sample presented some severe cost implications. Households with someone 15-35 only accounted for approximately 32% of the households in Baltimore. Therefore, two out of every three households found were not eligible for the study. Since screening households is difficult, time consuming, and costly, it was determined to look for a sample design that might add more efficiency. Without doing so, survey costs could grow beyond tolerable levels.

Therefore, six months into the study, a more cost effective sample design was instituted. A dual frame stratified sample design was adopted which made use of a list sample approach combined with an RDD approach. With this approach, four sample strata were created:

1) Households in Baltimore identified on a list maintained by the Marketing Systems Group (MSG) as likely to have someone 15-35 years of age.
2) Households in Baltimore identified on a list maintained by MSG in which the ages of household residents were unknown.
3) Households in Baltimore identified on a list maintained by MSG as <u>not</u> likely to have someone aged 15-35.
4) All telephone numbers in Baltimore from the RDD sample frame that are <u>not</u> on any of the three above lists.

Creating the strata in this manner produced two important results. First, four non-overlapping strata were created. A household could appear in one and only one stratum with the strata defined this way. Second, efficiencies could be gained by utilizing the lists to reduce household screening. The net result would hopefully be a gain in efficiency and a reduction in survey costs. A thorough discussion of the efficiencies gained by switching to this new design can be found in Roman, et. al. [1].

The data collection was designed in the following manner. An advance letter was sent to all sample telephone numbers for which a matching address could be obtained. Within a week of the expected receipt of this letter, interviewers in the Boston telephone facility at CSR would call the sample telephone numbers. The residential status of the telephone number would be ascertained and if a residential household was located, attempts were made to screen the household for eligibility. Eligibility criteria were:

1) The household was located within the city of Baltimore.
2) There was at least one person between the ages of 15 and 35 currently living in the household.
3) The selected eligible respondent spoke English.
4) The household had a touch-tone telephone.

If the household met all these criteria, the number of people aged 15-35 living in the household was ascertained. If more than one eligible person was living in the household, then one was randomly selected. If the selected person was an adult, then an attempt was made to interview the adult at that time. If the selected person was a minor (15-17 years old), then an attempt was made to get a parent or guardian on the phone to gain parental consent. No minor could be interviewed without parental consent. Once parental consent was obtained, an attempt was made to interview the minor.

To conduct the interview, the interviewer initially introduced the survey, the sponsors of the survey, stated that cooperation was voluntary, described the topics to be covered, and informed each respondent they would be paid $20. If the respondent agreed, the interviewer would establish a three-way telephone connection among the interviewer, the respondent and the TACASI computer located in RTI's telephone facility in North Carolina. The interviewer would stay on the line while the computer asked the first few non-sensitive questions. The interviewer would make sure that the respondent knew how to answers questions by hitting the touch tone keys on the telephone, that the connection was working properly and that the respondent was comfortable with the setting. At that point, the interviewer would tell the respondent they were hanging up, but would return when the interview was complete. The interviewer then hung up and allowed the respondent to answer questions in private to the computer. Once the interview was complete, the computer called the interviewer back and reestablished a three way connection. The computer then hung up and the interviewer obtained mailing address information in order to pay the respondent. The interviewer also enlisted cooperation in the biospecimen collection component of the study by explaining the use and mailing of the specimen kit and receipt of an additional $40 for sending in a sample.

Data Collection Results

Over the three year study period, 2936 interviews were completed. Of these, 2522 people (85.9%) agreed to accept a urine specimen kit in the mail. Of those that received the specimen kit, 2136 (84.7%) returned a sample to the lab. This meant that specimens were received for testing from 72.8% of all people interviewed. Rates of specimen receipt were very consistent across the four strata used in the sample design.

Table 1 displays residential rates and age eligibility rates among the original RDD design and the two primary list strata. It is easily seen that the list strata were far more efficient in terms of identifying households with age eligible respondents.

Table 1: Comparison of eligibility rates

Sample Source	Residential Rates	Age Eligibility Rates	Overall Eligibility Rates
Original RDD Sample	30.8%	29.9%	9.2%
List with someone 15-35	73.0	55.0	40.2
List with ages unknown	55.3	35.0	19.4

In terms of sampling efficiency, the stratum defined by list households with ages of residents unknown was over twice as efficient in finding households with an eligible respondent as compared to the original RDD sample (19.4% vs. 9.2%). For the stratum with listed households that were expected to have someone 15-35 years of age, this stratum was over four times as efficient in finding eligible households compared to the original RDD sample (40.2% vs. 9.2%). The list had the desired effect of producing a far more efficient method of locating eligible households.

The initial question of interest then becomes what bias may exist in critical survey measures derived from increased, or even exclusive, reliance on the lists. Secondary questions involve what estimated cost savings are derived and what potential increases in variance estimates are obtained from increased use of these lists.

Examination of Bias

Since the first year of data collection was a transition year in which the sample design was switched from the original RDD to the stratified dual frame approach, the analyses cited in this paper rely exclusively on the second and third years of data collection. In these years, the sample design was strongly entrenched and interviewers were more experienced. This provided for a more stable period to evaluate the potential effects of the list approach.

Since estimated prevalence rates for STIs were the most critical survey measures, it was decided to examine these for potential bias issues. In particular, rates for Chlamydia and trichomoniasis were evaluated. Rates of gonorrhea were not considered as the rates were so low as to make it impossible to conduct any comparative analysis. To begin the analysis, rates were constructed for the combined list strata and compared to the stratum consisting of the remaining RDD telephone numbers not on any list. Table 2 presents these results.

Table 2: Estimated prevalence rates

	List Strata	Remaining RDD Stratum	Overall
Chlamydia	2.30%	5.14%*	3.56%
Trichomoniasis	7.33	9.15	8.14

* implies difference is significant at the 0.05 level.

These results are not completely unexpected as characteristics of people found on lists can frequently differ from those not on lists. These differences in turn may explain why estimates themselves differ. Table 3 examines differences in strata composition by gender, race, age and education, in comparison to estimates from the 2006 American Community Survey (ACS).

Table 3: Comparison of Demographic Characteristics

	List 15-35:	List Age Unknown:	Combined Lists:	RDD with Lists Removed:	2006 ACS:
Gender:					
Male	37.4%	38.5%	38.3%	38.4%	52.2%
Female	62.6	61.5	61.7	61.6	47.8
Race:					
Black	50.9	60.4	58.6	83.7	62.3
Non-black	49.1	39.6	41.4	16.3	37.7
Age:					
15-17	16.7	15.3	15.6	19.2	14.8
18-21	15.7	16.4	16.2	23.8	21.6
22-24	12.6	15.4	14.9	15.1	14.4
25-30	23.3	33.7	31.6	25.7	26.7
31-35	31.7	19.2	21.7	16.2	22.4
Education:					
Aged 25-35:					
< high school	2.8	12.5	10.5	16.7	11.3
High school[1]	46.2	35.4	37.6	59.2	44.8
4YR college +	51.0	52.1	51.9	24.1	43.9
Aged 18-24:					
< high school	7.7	30.2	26.1	16.9	20.2
High school[1]	82.3	53.2	58.5	80.1	69.0
4YR college +	10.0	16.6	15.4	3.0	10.8

[1] This row includes people with a high school diploma, a G.E.D. or some college beyond high school but less than a 4-year college degree.

This table shows that differences exist between the list strata and the remaining RDD stratum for race, age and education while gender is quite consistent across the strata.

To examine whether these demographic characteristics explain the differences in prevalence rates seen in Table 2, data from the list strata were taken and post stratification adjustments to the 2006 ACS were applied for the age, gender, race and education categories displayed in Table 3. These were the same categories used in sampling weight adjustments applied to the survey data as a whole. After adjusting these list strata data, the recomputed prevalence rates are shown in table 4.

Table 4: Recomputed prevalence rates (after post-stratification weight adjustments)

	Overall Study Estimate	List Strata Alone and re-weighted
Chlamydia	3.56%	2.55%*
Trichomoniasis	8.14	7.94

* implies difference is significant at the 0.05 level.

The results of this exercise were quite surprising. They show that an estimated prevalence rate for trichomoniasis computed only from the list strata was not biased at all and in fact was extremely close to the overall study estimate computed from all available data after post stratification weight adjustments were applied. This means the lists alone could have been used to produce an estimate of this rate with the maximum amount of cost savings this implies. The observed differences in this prevalence rate across strata are sufficiently explained by the differences in demographic composition of the strata and an adjustment can correct for this.

However, this is not the case for the prevalence rate for chlamydial infection. In this instance, there is still a significant difference between the estimate produced from the lists alone and the estimate produced from all available data. Post stratification adjustments could not account for the observed original differences across strata. In fact, adding marital status and employment status to the post stratification cells, did not change the observed estimate. The difference in chlamydial infection between strata cannot be explained by controlling for simple demographic differences.

It is difficult to surmise what might cause prevalence rates for one STI to be biased using lists alone and yet unbiased for another STI. A far more extensive analysis would be required to try to discover other factors that may be a source of this effect. In fact, it may be impossible to ever truly know. This demonstrates the great difficulty in trying to determine in advance what estimates may or may not be affected by using a sample frame that is less inclusive. It brings home the point that researchers proceed at their own peril once this road is travelled.

Cost and Variance Estimates

The sample design used in this particular study was quite conservative. 24% of all sample cases were taken from the list stratum with people expected to be 15-35. Another 48% was taken from the list stratum with unknown ages and the remaining 28% was taken from the remaining RDD stratum. Less than 1% was taken from the list stratum not expected to have anyone in the 15-35 age range as the efficiency of sampling within this stratum was extremely low. This particular allocation resulted in an 11% cost savings over the original RDD design. It also resulted in estimated standard errors that had a design effect of 1.25. Therefore, although gains were made for cost (or estimated sample size for a fixed cost), there was a loss in precision due to the stratified sample allocation. This is not unexpected or unusual. We estimate that other sample allocations across strata could have increased the cost savings to about 20%. Obviously, going with lists alone would have increased the cost savings even more and probably up to about 30% but this would have resulted in some observed biases. In the future, work needs to be done in developing an optimum allocation for estimates such as these.

Summary

In an age with declining response rates to telephone surveys and increased loss of population coverage for telephone sample frames due to cell-phone only households, there is a great rush to move to more efficient sample frames and modes of data collection. These include list frames such as those used in this study as well as internet panels that may use web surveys for data collection. Many other alternatives also exist in practice. These alternate frames usually share a trait in common in that they either lack population coverage or have unknown biases from frame construction.

The example presented in this paper is not meant to disparage use of such frames, rather to illustrate the need to evaluate frames in terms of study goals and to demonstrate the difficulty in determining when a bias may be present and when it may not. In this exercise, use of list frames would have led to biases in estimation of one STI but not the other. Whether and how other STI estimates would be affected is not known. If this is the case for one study outcome, STIs, how does one evaluate potential biases in other survey measures? This is a very troubling question. As a lesson learned or re-learned, it is always important to assume little when designing studies and to utilize sample frames with high levels of population coverage, ones with known probabilities of how the frame was developed and with data collection protocols that lead to as high response rates as possible. Cost containment is always critical, but accuracy should never be sacrificed. Assuming biases may not exist is a very dangerous path.

References

1. Roman, Anthony M.; Elizabeth Eggleston; Charles F. Turner; Susan M. Rogers; Rebecca Crow; Sylvia Tan; James R. Chromy; Laxminarayana Ganapathi. 2008. “Effects of Sampling and Screening Strategies in an RDD Survey”, In JSM Proceedings, Survey Research Methods Section. Alexandria, VA: American Statistical Association, pp 2712-2719.