

**TECHNICAL PAPERS ON  
HEALTH AND BEHAVIOR MEASUREMENT**

TECHNICAL PAPER 86

**Monitoring STIs Survey Program (MSSP)**

**Methodological Report**

Anthony M. Roman  
Center for Survey Research  
University of Massachusetts, Boston

**Reference Citation**

A.M. Roman (2010). Monitoring STIs Survey Program (MSSP): Methodological Report. Technical Papers on Health and Behavior Measurement, No. 86. Washington DC: Program on Health and Behavior Measurement, Research Triangle Institute, 2010.

# **Monitoring STIs Survey Program (MSSP)**

## **Methodological Report**

### **Prepared by:**

Anthony M. Roman  
Senior Statistician  
Center for Survey Research  
University of Massachusetts-Boston

### **For:**

Susan M. Rogers  
Elizabeth Eggleston  
Principal Investigators  
Research Triangle Institute

and

Charles F. Turner  
Principal Investigator  
City University of New York  
Queens College and Graduate Center

February 2010

## Table of Contents

	<u>Page</u>
I. Background .....	3
II. Sample Design and Other Methodological Issues .....	4
III. Data Collection Procedures .....	7
IV. Data Collection Results .....	9
A. Screening Households .....	9
B. Interviewing Households .....	14
C. Specimen Cups .....	20
D. Parental Consent .....	23
E. Vaginal Swabs .....	23
V. Weighting .....	24
VI. Variance Estimation .....	25
VII. Survey Effort .....	27
VIII. Conclusions .....	28
IX. Acknowledgements .....	29
References .....	30

## Tables

	<u>Page</u>
1. Screening Results .....	10
2. Interviewing Results .....	16
3. Acceptance and Return of Specimen Kits by Year .....	20
4. Acceptance and Return of Specimen Kits by Stratum .....	21
5. Acceptance and Return of Specimen Kits by Adult/Teen .....	22
6. Unweighted and Weighted Sample Totals by Year and Stratum .....	27
7. Effort Expended to Complete Interviews .....	28

## I. 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 the interview. Those who completed an interview were asked to participate further by submitting a biospecimen that would be tested for three specific sexually transmitted infections (STIs), namely gonorrhea, Chlamydia, and trichomonas. For those who agreed, a specimen collection kit was mailed to the respondent along with instructions on use of the kit, and 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 this 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 Chlamydia as is required by Maryland law. For any selected respondent under 18 years of age, parental permission was first obtained. This was done over the telephone and done separately for permission to interview the teen and then permission to offer them the urine specimen cup. If permission was obtained, then attempts were made to contact the teen and complete the interview. For teens that tested positive, only the teen was notified of the test result, not the parent.

CSR was responsible for developing the sample design for this survey and for purchasing the sample. CSR was also responsible for screening all sample telephone numbers for eligibility, randomly selecting an eligible household respondent, getting parental consent when required, enlisting cooperation, and attempting to get respondents to accept the urine specimen kit. CSR was also responsible for mailing advance letters to sample households, for mailing specimen kits to those who accepted them, and for mailing all respondent pay. CSR also attempted to call all those who did not follow through on mailing kits to the lab and attempting to get them to do so. RTI was responsible for questionnaire development, overall management of all aspects of the survey, coordination of all participating organizations, programming and maintenance of the computers required for conducting all TACASI interviews, creating all data files for analysis, performing all data analysis and for disseminating results. UNC was responsible for all testing of biospecimens (urine and vaginal swabs) for the three specified STIs. JHU was responsible for

interactions with the BCHD and all community efforts within the city of Baltimore. This survey was truly a coordinated effort attempting to obtain the best results possible.

This report is not an analytic report of results. It is documentation of the methodological design and the results from data collection efforts which were the responsibility of CSR. Results from data analyses will be reported in other papers.

## **II. Sample Design and Other Methodological Issues**

Original Sample Design. The original sample design called for a pure random-digit-dialed (RDD) sample of all telephone numbers serving the city of Baltimore, Maryland. Furthermore, the survey would be conducted over three years by collecting data quarterly. Therefore, 12 separate quarterly samples would be required. After data collection began, it quickly became apparent that this design had two serious drawbacks. They were:

- 1) Quarterly samples would be difficult to manage and very costly, and
- 2) A pure RDD sample approach would also be very costly.

Quarterly samples had two issues related to them. The first was that the target age group was people 15-35 years of age. This particular age group is very difficult to track down and get on the phone. It takes many call attempts and substantial time. Having three month set intervals to conclude all interviews and get all specimen cups returned works against the goals of trying to maximize survey response rates and maximize the return of urine specimen cups. The second issue was that the design proved not to be cost effective. In order to attempt to complete the interviews in three month windows, a large sample needed to be released early and then worked. This led to chasing difficult cases and only difficult cases for an extended period of time in month three of the quarter. This would then be replicated 12 times across the 12 quarterly data collections. The most costly time in any survey occurs when the survey is being closed out and there are smaller numbers of cases to work. There is just a lack of efficiency present when dealing with this situation. The quarterly design implied there would be 12 closing out periods. With the added costs and the projected lower response rates, 12 quarterly samples seemed to be too inefficient a design.

In addition, the pure RDD sample also had cost implications. Households with someone 15-35 only account 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.

Revisions to Original Design. Therefore, six months into the study, a video conference call was held among all interested parties at RTI and CSR. The purpose was to discuss a more cost effective sample design. CSR led this discussion inviting Jack Fowler to participate, as he had a high degree of experience to call upon. Several important changes to the study came out of this video conference. Regarding the sample design, a totally new design was recommended and implemented. First, the quarterly samples were dropped in favor of yearly samples. This allowed for a more appropriate time to pursue interviews with difficult sample households. In addition, 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 not likely to have someone aged 15-35.
- 4) All telephone numbers in Baltimore from the RDD sample frame that are not on any of the three lists just detailed.

Creating the strata in this manner produced two important results. The first was that four non-overlapping strata were created. A household could appear in one and only one stratum with the strata defined this way. The second was that 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. <sup>1</sup>.

A second methodological change adopted during this video conference related to within household respondent selection. Originally, each eligible household member was given an equal chance to be selected as the participant in this survey. Since for this age group, getting people on the telephone is a time consuming and costly task, a new respondent selection model was adopted which gave an increased probability to selecting the person on the telephone answering the screening questions if that person was an eligible respondent. The procedure worked in the following manner. If the person answering the screening questions

was not eligible, then as before, all eligible people in the household were given an equal chance of selection, namely  $(1/n)$  if  $n$  was the number of eligible people in the household. If the person answering the screening questions was found to be eligible, then that person was assigned a  $(2/(n+1))$  probability of selection with all other eligible people in the household getting probabilities of  $(1/(n+1))$  of being selected. This change was made to increase efficiency, reduce cost and still maintain a probabilistic method of within household respondent selection. Again, a complete description of this change and the effects of this change are found in Roman, et.al.<sup>1</sup>.

Two additional methodological changes were adopted during this video conference. They were:

- 1) Increasing the respondent pay for people who failed to return their specimen after repeated reminders from \$40 to \$100. Since obtaining specimens for STI testing was a critical aspect of the study, it was determined to be worth the added cost to get more specimens returned. Again refer to Roman et. al.<sup>1</sup> for more details on the effects of this change.
- 2) To offer an option to people to have someone come to their home to pick up the specimen cup if they preferred. This again was done to try to increase the number of specimens returned.

One additional methodological change should be mentioned for completeness. This involved the screening questions targeting households with someone 15-35 years of age. When the survey began, a single question was asked. This question was:

How many people aged 15 through 35 currently live in this household?

From early survey results, it was found that only about 21% of all households were being identified as eligible with this approach. This was significantly lower than the expected 32% based upon recent estimates from the American Community Survey<sup>2</sup>. The screening questions were altered to ask the following two questions in the order they are given:

How many people aged 36 or older currently live in this household?

How many people aged 15 through 35 currently live in this household?

With the added question, about 31% of all households were found eligible. This rate was much closer to the expected rate of 32%. This trend held up throughout the entire data collection period.

After these changes, data collection proceeded for the remainder of the study with only one additional methodological change. Approximately 26 months into the data collection, it was decided to ask women to provide both a vaginal swab and a urine specimen since the swab was considered a more accurate test. This was implemented over the remaining months of data collection.

### **III. Data Collection Procedures**

The data collection was designed in the following manner. An advance letter was sent to all sample telephone numbers for which an address was known. 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) People in the household 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 using procedures outlined previously. 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 teen (15-17 years old), then an attempt was made to get a parent or guardian on the phone to gain parental consent. No teen could be interviewed without parental consent. If parental consent was obtained, an attempt was made to interview the teen.

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 answer 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. The interviewer then obtained mailing address information in order to pay the respondent. The interviewer also enlisted cooperation by telling the respondent about receiving the specimen kit and that they would be paid an additional \$40 for sending in a sample.

An initial interviewer briefing was held on August 10, 2006 and telephone interviewing began on August 11, 2006. It quickly became apparent that a serious problem existed in the connection to the TACASI computer. Background music could be heard and connections were not solid. Interviewing was immediately stopped. The problems came as a complete surprise because the entire system was tested a few months prior without any difficulty. It was found that a new telephone system was installed at RTI in the interim and although no one had any reason to expect complications, they did occur. The problems took awhile to solve. Therefore interviewing did not begin until September 5, 2006, just after Labor Day. Interviewers were re-briefed on that day as well. No other connection problems occurred.

For respondents accepting urine specimen cups, kits were mailed by CSR no longer than three days after completion of the interview. If CSR did not obtain verification that the lab received the kit within three weeks of the initial mailing, then CSR began calling the respondent to urge them to send in the kit. These calls could go on for several months if necessary as getting returned specimens was critical to the objectives of the study. As previously detailed, the respondent pay was increased from \$40 to \$100 for those difficult to reach respondents.

A total of 33 CSR interviewers worked on the study over the three year period. Since the survey topic was sensitive, the age group difficult to contact, the interviews all from an urban center, and parental consent was involved, only the most experienced and successful interviewers could remain on the study. This was done in an attempt to maximize both response rates and specimen return rates. During the three year data collection period, five meetings among interviewers were held. The purpose of these meetings was for interviewers to be able to exchange tips on how they enlist cooperation and get people to complete the interview. The purpose was also to reenergize the staff and keep them current on the progress of the study.

Interviewing continued until June 14, 2009 and calls to get people to return specimens continued until August 15, 2009.

#### **IV. Data Collection Results**

##### **A. Screening Households**

The data collection effort begins with having to screen households for eligibility over the telephone. This has become a more difficult undertaking since 2000 because of the prevalence of voice mail systems, caller ID systems, people who use their land line telephones less since they have cell phones and a number of other reasons. Table 1 displays the results for every telephone number dialed across the three year period. The table displays results for each individual year as well as for the combined three year period. It also displays results within each year by sample stratum.

The table shows that screening rates hovered around 70% for the three year period (70.68% in year 1, 69.54% in year 2 and 64.58% in year 3). The reason that year 3 is a bit lower is due to the culmination of data collection in June. This allowed for less time to pursue the more difficult cases to get a positive resolution. The lengths taken to pursue each telephone number will be described later in the section on survey effort. Overall, the screening rate across the three year period was 68.84%. As a point of comparison, a similar study using TACASI conducted in Baltimore in 1999 produced a screening rate of 80.23%<sup>3</sup>. This demonstrates how much more difficult screening households by telephone has become over the last decade.

Table 1 also shows that screening rates were fairly consistent across survey stratum. The strata that were formed from the list samples did just slightly better than their RDD counterpart (69.6% for the list from households with someone expected to be 15-35, 69.8% for the list from households with unknown ages, 73.6% for the list with from households not expected to have anyone 15-35 as compared to 65.2% for the RDD sample after the lists were removed). It is not that surprising that the list stratum from households not expected to have someone 15-35 has the highest screening rate as these households would include more elderly people who would be expected to be easier to contact and easier to screen. The fact that the other strata vary only slightly is also not that surprising as all households are difficult to screen, regardless if they end up on a list or not.

**Table 1: Results from Screening Sample Households by Stratum**

**Year 1:**

Stratum	Complete Screeners	Nonresidential	No Contact <sup>1</sup>	Age Ineligible	Other Ineligible <sup>2</sup>	Refusals	Limits <sup>3</sup>	Total Cases	Resolution Rate <sup>4</sup>	Eligibility Rate <sup>5</sup>	Screening Rate <sup>6</sup>
List with 15-35	273	180	16	169	9	98	91	836	98.1%	60.5%	70.36%
List no one 15-35	9	27	1	40	1	10	12	100	99.0	18.0	69.44
List Age Unknown	552	1218	129	794	96	293	258	3340	96.1	38.3	72.17
RDD after stratifying	400	8397	877	721	51	263	259	10968	92.0	34.1	67.78
RDD before stratifying	942	9858	1127	2116	98	665	567	15373	92.7	29.9	71.19
Total	2176	19680	2150	3840	255	1329	1187	30617	93.0	34.7	70.68

<sup>1</sup> Telephone numbers with no contact of any kind including answering machines after at least two weeks of calling and 15 attempts.

<sup>2</sup> These include households with no one who spoke English, those with no touch-tone phone, those not in Baltimore, respondents too ill to respond and other such conditions.

<sup>3</sup> telephone numbers believed to be residential which would not respond after numerous call attempts over at least three weeks.

<sup>4</sup> The rate at which telephone numbers were determined to be residential or not  $(\text{Total Cases} - \text{No Contact}) / (\text{Total Cases})$ .

<sup>5</sup> The rate at which residential households were confirmed to be eligible for the study.

$$(\text{Complete Screeners}) / (\text{Complete Screeners} + \text{Age Ineligible} + \text{Other Ineligible})$$

<sup>6</sup> The rate at which telephone numbers were successfully screened.

$$(\text{Complete Screeners} + \text{Age Ineligible} + \text{Other Ineligible}) /$$

$$(\text{Complete Screeners} + \text{Age Ineligible} + \text{Other Ineligible} + \text{Refusals} + \text{Limits} + (0.04)(\text{No Contacts}))$$

The rate of 0.04 was determined as the average rate that No Contacts proved to be residential by CSR over many studies where samples of No Contacts are followed up with telephone business offices.

**Year 2:**

Stratum	Complete Screeners	Nonresidential	No Contact <sup>1</sup>	Age Ineligible	Other Ineligible <sup>2</sup>	Refu-sals	Limits <sup>3</sup>	Total Cases	Resolution Rate <sup>4</sup>	Eligibility Rate <sup>5</sup>	Screening Rate <sup>6</sup>
List with 15-35	411	308	40	306	37	128	152	1382	97.1%	54.5%	72.78%
List no one 15-35	16	20	2	45	0	8	6	97	97.9	26.2	81.33
List Age Unknown	752	2228	185	1329	165	405	469	5533	96.7	33.5	71.83
RDD after stratifying	523	14336	1318	1060	106	366	485	18194	92.8	31.0	65.14
Total	1702	16892	1545	2740	308	907	1112	25206	93.9	35.8	69.54

<sup>1</sup> Telephone numbers with no contact of any kind including answering machines after at least two weeks of calling and 15 attempts.

<sup>2</sup> These include households with no one who spoke English, those with no touch-tone phone, those not in Baltimore, respondents too ill to respond and other such conditions.

<sup>3</sup> telephone numbers believed to be residential which would not respond after numerous call attempts over at least three weeks.

<sup>4</sup> The rate at which telephone numbers were determined to be residential or not (Total Cases – No Contact)/(Total Cases).

<sup>5</sup> The rate at which residential households were confirmed to be eligible for the study.

(Complete Screeners)/(Complete Screeners + Age Ineligible + Other Ineligible)

<sup>6</sup> The rate at which telephone numbers were successfully screened.

(Complete Screeners + Age Ineligible + Other Ineligible)/

(Complete Screeners + Age Ineligible + Other Ineligible + Refusals + Limits + (0.04)(No Contacts))

The rate of 0.04 was determined as the average rate that No Contacts proved to be residential by CSR over many studies where samples of No Contacts are followed up with telephone business offices.

**Year 3:**

Stratum	Complete Screeners	Nonresidential	No Contact <sup>1</sup>	Age Ineligible	Other Ineligible <sup>2</sup>	Refu-sals	Limits <sup>3</sup>	Total Cases	Resolution Rate <sup>4</sup>	Eligibility Rate <sup>5</sup>	Screening Rate <sup>6</sup>
List with 15-35	310	471	28	276	15	171	145	1416	98.0%	51.6%	65.47%
List no one 15-35	2	38	6	35	0	5	12	98	93.9	5.4	68.51
List Age Unknown	590	2818	212	1037	100	533	365	5655	96.3	34.2	65.59
RDD after stratifying	218	8237	806	558	37	267	203	10326	92.2	26.8	61.83
Total	1120	11564	1052	1906	152	976	725	17495	94.0	35.2	64.58

<sup>1</sup> Telephone numbers with no contact of any kind including answering machines after at least two weeks of calling and 15 attempts.

<sup>2</sup> These include households with no one who spoke English, those with no touch-tone phone, those not in Baltimore, respondents too ill to respond and other such conditions.

<sup>3</sup> telephone numbers believed to be residential which would not respond after numerous call attempts over at least three weeks.

<sup>4</sup> The rate at which telephone numbers were determined to be residential or not  $(\text{Total Cases} - \text{No Contact}) / (\text{Total Cases})$ .

<sup>5</sup> The rate at which residential households were confirmed to be eligible for the study.

$(\text{Complete Screeners}) / (\text{Complete Screeners} + \text{Age Ineligible} + \text{Other Ineligible})$

<sup>6</sup> The rate at which telephone numbers were successfully screened.

$(\text{Complete Screeners} + \text{Age Ineligible} + \text{Other Ineligible}) /$

$(\text{Complete Screeners} + \text{Age Ineligible} + \text{Other Ineligible} + \text{Refusals} + \text{Limits} + (0.04)(\text{No Contacts}))$

The rate of 0.04 was determined as the average rate that No Contacts proved to be residential by CSR over many studies where samples of No Contacts are followed up with telephone business offices.

**Total (Combined Years 1 through 3):**

Stratum	Complete Screeners	Nonresidential	No Contact <sup>1</sup>	Age Ineligible	Other Ineligible <sup>2</sup>	Refusals	Limits <sup>3</sup>	Total Cases	Resolution Rate <sup>4</sup>	Eligibility Rate <sup>5</sup>	Screening Rate <sup>6</sup>
List with 15-35	994	959	84	751	61	397	388	3634	97.7%	55.0%	69.62%
List no one 15-35	27	85	9	120	1	23	30	295	97.0	18.2	73.63
List Age Unknown	1894	6264	526	3160	361	1231	1092	14528	96.4	35.0	69.79
RDD after stratifying	1141	30970	3001	2339	194	896	947	39448	92.4	31.1	65.18
RDD before stratifying	942	9858	1127	2116	98	665	567	15373	92.7	29.9	71.19
Total	4998	48136	4747	8486	715	3212	3024	73318	93.5	35.2	68.84

<sup>1</sup> Telephone numbers with no contact of any kind including answering machines after at least two weeks of calling and 15 attempts.

<sup>2</sup> These include households with no one who spoke English, those with no touch-tone phone, those not in Baltimore, respondents too ill to respond and other such conditions.

<sup>3</sup> telephone numbers believed to be residential which would not respond after numerous call attempts over at least three weeks.

<sup>4</sup> The rate at which telephone numbers were determined to be residential or not  $(\text{Total Cases} - \text{No Contact}) / (\text{Total Cases})$ .

<sup>5</sup> The rate at which residential households were confirmed to be eligible for the study.

$(\text{Complete Screeners}) / (\text{Complete Screeners} + \text{Age Ineligible} + \text{Other Ineligible})$

<sup>6</sup> The rate at which telephone numbers were successfully screened.

$(\text{Complete Screeners} + \text{Age Ineligible} + \text{Other Ineligible}) /$

$(\text{Complete Screeners} + \text{Age Ineligible} + \text{Other Ineligible} + \text{Refusals} + \text{Limits} + (0.04)(\text{No Contacts}))$

The rate of 0.04 was determined as the average rate that No Contacts proved to be residential by CSR over many studies where samples of No Contacts are followed up with telephone business offices.

A few other results from Table 1 should be mentioned. The resolution rate is the rate for which telephone numbers could be definitely classified as either connecting to a residential household or not. For example, a telephone number that was always an answering machine would be classified as residential unless the answering machine message clearly indicated a business. Either way, the telephone number was resolved. Telephone numbers which could not be resolved were those that were dialed many times with only rings but no other outcome, including answering machines, voicemail, recorded messages or someone picking up the telephone. Overall, 93.5% of all telephone numbers dialed were definitely resolved as residential or not. This is a very high rate. CSR attempts to hit rates of 90% or higher on all telephone studies. The previously referenced 1999 study in Baltimore had a resolution rate of 92.0%<sup>3</sup>. Therefore, the rate over the 2006-2009 period of this study was actually a bit higher than in 1999. The effort extended to produce this high of a rate is rewarded by the fact that over time, the percentage of unresolved telephone numbers from CSR telephone studies that prove to be residential has consistently been 4%. This estimate is obtained by taking samples of unresolved telephone numbers and pursuing their residential status with telephone business offices. This rate is applied to the unresolved telephone numbers for AAPOR response rate calculations.

Another result that should be mentioned is the eligibility rates obtained by stratum. Of the households screened, 55.0% of those from the list stratum expected to have someone 15-35 were eligible for the study. This is lower than hoped, but still a significant gain as will be seen when compared to the other strata. For the list stratum with ages of residents unknown, the eligibility rate was 35.0% while for the list stratum of households not expected to have anyone 15-35 had an eligibility rate of only 18.2%. The list strata did offer gains in efficiencies. The RDD stratum had rates of only about 30%. Therefore, use of the lists did accomplish a gain in sampling efficiency. How this manifested itself in terms of cost savings is discussed in Roman, et. al.<sup>1</sup>.

## **B. Interviewing Households**

The other part of conducting successful telephone data collections is being able to complete interviews once an eligible household has been identified. For this particular study, this is not as easy as it may appear. The study involves an urban center which is predominantly African-American and is lower in income. The age group targeted is people 15-35 years old. These are all younger people who would be expected to spend less time at home. For 15-17 year olds, parental consent is also required.

Table 2 presents the results for interviewing eligible households. It is again divided up by the three years of data collection and by stratum within year. A summary is also included for all three years combined. Once again, rates are fairly consistent across years and hover around 60% (57.35% in year 1, 60.69% in year 2 and 58.48% in year 3). Year 3 may have been over 60% if the same amount of time to follow-up difficult respondents could have been allowed. Overall, 58.48% of all eligible households were successfully interviewed. This can be compared to the 1999 Baltimore study in which 62.29%<sup>3</sup> of all eligible households were interviewed. This rate is quite comparable over the nearly 10 year span showing that it is primarily screening households that causes the drop in overall survey response rates. In addition, since the 1999 study targeted 18-45 year olds, it would be expected to be a bit higher since parental consent is never required and 36-45 year olds would be expected to be a bit easier to interview.

When comparing interviewing rates across strata, there is not much variation. Overall, the list from households expected to have someone 15-35 was successfully interviewed at a rate of 58.6%. This compares to the list with unknown ages of residents at 61.3% and the RDD sample at 57.7%. The list from households which were not expected to have anyone 15-35 produced the lowest interview rate at 51.9%. However, this was based on only 27 total cases, so not much can be made of that number. Overall, it can be stated that once households are screened, the rate at successfully interviewing a respondent was not different across strata.

Table 2 also displays final survey response rates. These rates are AAPOR Method 3 calculated response rates. Overall, the study had a 40.44% response rate over the three year study period. Year 3 had the lowest response rate at 37.77%, but again that was probably lower due to less amount of time to follow the very difficult cases. Whereas, higher response rates are always preferred, surpassing 40% at a time of declining response rates and for a study that involved an urban area, a high minority population, a lower income population, a sensitive topic, and some parental consent, this must be considered successful. The primary reason for lower response rates was shown to be the extreme difficulty in screening households.



**Table 2: Results from Interviewing Successfully Screened Households by Stratum**

**Year 1:**

Stratum	Complete Interviews	Refusals	Noninterview-Limits <sup>1</sup>	Total Cases	Interview Rate <sup>2</sup>	Response Rate <sup>3</sup>
List with 15-35	151	89	33	273	55.31%	39.15%
List no one 15-35	6	1	2	9	66.67	46.30
List Age Unknown	352	151	49	552	63.77	46.02
RDD after stratifying	220	130	50	400	55.00	37.35
RDD before stratifying	519	290	133	942	55.10	39.23
Total	1248	661	267	2176	57.35	40.53

<sup>1</sup> These are cases declared a non-interview after many call attempts with no interview completed

<sup>2</sup> This is the rate at which screened households were successfully interviewed.

(Complete Interviews)/(Complete Interviews + Refusals + Noninterview-Limits)

<sup>3</sup> This is the product of the screening rate and the interview rate (screening rate \* interview rate)

It is equivalent to a response rate computed using AAPOR Method 3.

**Year 2:**

Stratum	Complete Interviews	Refusals	Noninterview-Limits <sup>1</sup>	Total Cases	Interview Rate <sup>2</sup>	Response Rate <sup>3</sup>
List with 15-35	242	126	43	411	58.88%	42.85%
List no one 15-35	7	8	1	16	43.75	35.58
List Age Unknown	464	232	56	752	61.70	44.32
RDD after stratifying	320	152	51	523	61.19	39.86
Total	1033	518	151	1702	60.69	42.20

<sup>1</sup> These are cases declared a non-interview after many call attempts with no interview completed

<sup>2</sup> This is the rate at which screened households were successfully interviewed.

(Complete Interviews)/(Complete Interviews + Refusals + Noninterview-Limits)

<sup>3</sup> This is the product of the screening rate and the interview rate (screening rate \* interview rate)

It is equivalent to a response rate computed using AAPOR Method 3.

**Year 3:**

Stratum	Complete Interviews	Refusals	Noninterview-Limits <sup>1</sup>	Total Cases	Interview Rate <sup>2</sup>	Response Rate <sup>3</sup>
List with 15-35	189	82	39	310	60.97%	39.92%
List no one 15-35	1	0	1	2	50.00	34.26
List Age Unknown	345	180	65	590	58.48	38.36
RDD after stratifying	120	67	31	218	55.05	34.04
Total	655	329	136	1120	58.48	37.77

<sup>1</sup> These are cases declared a non-interview after many call attempts with no interview completed

<sup>2</sup> This is the rate at which screened households were successfully interviewed.

(Complete Interviews)/(Complete Interviews + Refusals + Noninterview-Limits)

<sup>3</sup> This is the product of the screening rate and the interview rate (screening rate \* interview rate)

It is equivalent to a response rate computed using AAPOR Method 3.

**Total (Combined Years 1 through 3):**

Stratum	Complete Interviews	Refusals	Noninterview-Limits <sup>1</sup>	Total Cases	Interview Rate <sup>2</sup>	Response Rate <sup>3</sup>
List with 15-35	582	297	115	994	58.55%	40.76%
List no one 15-35	14	9	4	27	51.85	38.18
List Age Unknown	1161	563	170	1894	61.30	42.78
RDD after stratifying	660	349	132	1141	57.84	37.70
RDD before stratifying	519	290	133	942	55.10	39.23
Total	2936	1508	554	4998	58.74	40.44

<sup>1</sup> These are cases declared a non-interview after many call attempts with no interview completed

<sup>2</sup> This is the rate at which screened households were successfully interviewed.

(Complete Interviews)/(Complete Interviews + Refusals + Noninterview-Limits)

<sup>3</sup> This is the product of the screening rate and the interview rate (screening rate \* interview rate)

It is equivalent to a response rate computed using AAPOR Method 3.

### C. Specimen Cups

One of the novel aspects of this study was that a completed interview was not the end of the data collection process, but the mid-point. After the interview was complete, interviewers talked to the respondents about their willingness to accept a urine specimen collection kit in the mail, and to submit a sample to a lab at UNC to be analyzed for the three specific sexually transmitted infections. This was complicated even further by the fact that they were notified in advance that positive test results would be reported to the BCHD as required by law. Pilot surveys demonstrated that respondents would generally be very amenable to accepting the specimen kit and then generally follow through and mail the specimen to the lab. Any kits that were not returned required extended pursuit of the respondent to see if they would comply.

Table 3 displays the results for accepting specimen kits and then following through on sending the specimen to the lab. These results are presented by year of data collection.

**Table 3. Acceptance and Return of Specimen Kits by Year**

	Completed Interviews	Agreed to Accept Kit	Percentage Agreed	Returned Kit	Percentage Returned <sup>1</sup>	Overall Percentage Returned <sup>2</sup>
Year 1	1248	1067	85.5%	878	82.3%	70.4%
Year 2	1033	889	86.1%	778	87.5%	75.3%
Year 3	655	566	86.4%	480	84.8%	73.3%
Combined years 1-3	2936	2522	85.9%	2136	84.7%	72.8%

<sup>1</sup> This is the percent of those who agreed to accept the kit that actually then returned the specimen.

<sup>2</sup> This is computed as number returned divided by completed interviews.

This table shows that the percentage of people who agreed to accept a kit was very consistent over the three year period and was approximately 86%. This is the same rate observed in the pilot studies. It appears that this is the rate that can be expected with a strong interviewer team and using the approaches detailed in the data collection section of this report. In one earlier pilot study, all people who declined to accept the cup were contacted again about 1-2 weeks later to see if they would change their mind. All people re-contacted again refused to accept the kit. Basically, people who are not willing to do this are very strong in their beliefs.

Regarding the percentage of those that agreed to accept the kit who then returned a specimen, Year 2 had the highest rate of 87.5%. Year 1 was the lowest (82.3%) and that is probably because interviewers were still developing the skills and approaches to get the best results. Year 3 was a little lower than Year 2 primarily because the study ended and a subset of people could not be followed for as long a period of time to get them to return the cup. Overall, the expectation is that getting approximately 75% of all people interviewed to both accept and then return the kit is the best rate that can be obtained. It should be noted that 9.8% of those that accepted the cup required follow-up calls to get them to return the cup. On average, 6.98 follow-up calls were made to get people to return the cup with a maximum of 47 calls. This demonstrates that without significant follow-up efforts, the results just cited cannot be attained.

Table 4 displays these same results by sample stratum.

**Table 4. Acceptance and Return of Specimen Kits by Stratum**

	Completed Interviews	Agreed to Accept Kit	Percentage Agreed	Returned Kit	Percentage Returned <sup>1</sup>	Overall Percentage Returned <sup>2</sup>
List 15-35	582	487	83.7%	434	89.1%	74.6%
List No 15-35	14	13	92.9%	11	84.6%	78.6%
List – Age Unknown	1161	995	85.7%	860	86.4%	74.1%
RDD before Strat.	519	437	84.2%	342	78.3%	65.9%
RDD after Strat.	660	590	89.4%	489	82.9%	74.1%
Combined All strata	2936	2522	85.9%	2136	84.7%	72.8%

<sup>1</sup> This is the percent of those who agreed to accept the kit that actually then returned the specimen.

<sup>2</sup> This is computed as number returned divided by completed interviews.

When examining the overall rate of return, it is seen that little variation exists across strata and that rates of about 74% are the norm. The stratum for RDD before stratification had by far the lowest return rate (65.9%), but all these interviews were down very early in Year 1 before interviewers became more experienced and better at getting people to accept the cup and then following-up to make sure they returned it. The list stratum for households not expected to have anyone 15-35 years of age had the highest return rate (78.6%), but since this was based upon such a small sample size, not much weight should be put on this result.

Some numbers internal to Table 4 need mentioning. The stratum for households expected to have someone 15-35 years of age had the lowest rate of people agreeing to accept the cup (83.7%). However, it also had the highest rate of people who actually returned the cup (89.1%). It is very difficult to assess why this might be the case. Alternatively, the RDD stratum after stratification shows a reverse tendency. It has the next to highest rate of people accepting the cup (89.4%) and the lowest rate of people actually returning the cup (82.9%). Again, it is difficult to assess why this occurred.

Table 5 displays these same results again, but this time broken out by adults (people 18 years old or older) and teen respondents (people 15-17 years of age).

**Table 5. Acceptance and Return of Specimen Kits by Adult/Teen**

	Completed Interviews	Agreed to Accept Kit	Percentage Agreed	Returned Kit	Percentage Returned <sup>1</sup>	Overall Percentage Returned <sup>2</sup>
Adults (18+)	2471	2114	85.6%	1774	83.9%	71.8%
Teens (15-17) <sup>3</sup>	465	408	87.7%	362	88.7%	77.9%
Everyone (All Ages)	2936	2522	85.9%	2136	84.7%	72.8%

<sup>1</sup> This is the percent of those who agreed to accept the kit that actually then returned the specimen.

<sup>2</sup> This is computed as number returned divided by completed interviews.

<sup>3</sup> The 465 completed interviews are for all teen interviews. For 22 of these interviews, parental consent was not given to offer the teen the specimen kit. If these are not considered then the percentage agreed becomes 92.1% and the overall percentage returned becomes 81.7%. Likewise, for everyone these percentages become 86.5% and 73.3% respectively.

Here it becomes obvious that teens were both more willing to accept the cup and then again more likely to return it. This led to an overall return rate for teens, (77.9%) that was greater than the rate for adults (71.8%). It is not possible to determine the exact reason for this difference, but it may be that the monetary incentive was more effective among teenagers.

Overall, all these results display a certain consistency which leads to the conclusion that it is possible, with strong effort, to get an overall 75% return for cups from a sample of respondents.

#### **D. Parental Consent**

One important aspect of this study involved gaining parental consent to both interview teens between the ages of 15 and 17, and to be able to offer them the opportunity to receive a urine specimen collection kit in the mail. Considering the sensitive nature of the questions within this survey and the facts that for teens that test positive for STIs, the BCHD is notified of this result while the teen's parents are not, parental consent may be considered difficult to attain. Overall, the parents of 684 teens were asked for permission to allow their children to participate in the study. Of these, 620 (90.6%) agreed to allow their children to be interviewed. 577 (84.4%) agreed to allow both the interview and the offer of the specimen kit. Therefore, a total of 43 parents agreed to the interview, but not the specimen testing. 64 parents (9.4%) refused to allow their children to participate at all. This was considered a great success.

#### **E. Vaginal Swabs**

During the third year of data collection, a decision was made to offer females vaginal swabs as well as urine specimen cups. The rationale was that swabs are considered more accurate, and having a set of respondents provide both a urine specimen and a swab would be a way to test for the accuracy of the urine specimen. Consideration was given to offering males urethral swabs, but this was dismissed due to cost considerations and the fact that it was felt that very few males would agree to accept such a swab and then provide a specimen.

During this period, the following results were obtained:



	<u>Teen</u>	<u>Adult</u>	<u>Total</u>
Females interviewed	10	82	92
Females agreed to accept urine			
Specimen kit	9	68	77
Females who provided urine specimen	9	53	62
Females who provided swab	6	40	46

From these results, it can be seen that although we did get females to send back vaginal swabs for testing, it was at a lower rate than for the urine specimens. 80.5% of all females who agreed to accept a urine specimen cup sent back a urine specimen to the lab (77.9% of adults and 100.0% of teens). However, only 59.7% of all females sent back a vaginal swab (75.5% of adults and 66.7% of teens). This was a much lower rate. Teens were better at returning swabs than adults, but both had lower rates than for urine specimens. It is possible that if only the swab was provided, that these percentages might be higher, as some females may have considered the urine sufficient for the study. We cannot tell this from these data. We can say that for future studies, getting vaginal swabs to be returned is quite possible but is likely to be a lower rate of return.

## **V. Weighting**

Weighting of survey data was done using a traditional approach. Weights were constructed by initially using the inverses of the probabilities of selection. These were modified by the following:

- 1) Number of land-line telephone numbers within the household. This adjustment was capped at 2. In other words, those households stating they had three or more land-line telephone numbers were treated as if they had two. This was done to not increase the variability among weights due to a small number of cases.
- 2) Within household selection of an eligible respondent. This was done taking into account the exact methodology used for selection and the exact probability of selecting the household respondent as described in the section on data collection procedures.
- 3) Household level survey nonresponse. This was computed separately for each stratum.

Finally, weights were post-stratified adjusted to four demographic characteristics. They were:

- 1) Age – in five groups (15-17, 18-21, 22-24, 25-30, 31-35)
- 2) Gender – Male and Female
- 3) Race – Black and non-black
- 4) Education – This first considered age by looking separately at 18-24 year olds and 25-35 year olds. Within each of these age groups, three cells were created as less than high school, high school and some college, and 4-year college degree or higher. Persons under 18 were not adjusted for education as virtually all have less than a high school education.

Also, large weights were trimmed to constrain individual cases from having large weights that could blow up variance estimates. Overall, about 4%-5% of weights were trimmed during this process. Final weights displayed an unequal weighting effect (UWE) of about 1.5 and a ratio of the maximum weight to the median weight of about 4.0.

It should be pointed out that weights were developed separately for each data collection year. In addition, since not only were interviews conducted, but respondents were also asked to send in a specimen, weights were developed separately for interviews and for the set of cases that had returned specimens. Weights were then placed on the final data sets. This included both interview and returned specimen weights. It also included weights before post stratification adjustment and after. This way, the effects of post stratification adjustment could be determined if desired.

## **VI. Variance Estimation**

Since stratification was employed, complex sampling methods must be used to create accurate estimates of survey variances and standard errors. The study used a unique id number for sample cases that was eight digits long. The seventh digit of this id number indicates the stratum. This digit is coded as follows:

- 0 – Random digit dialed stratum
- 1 – List stratum for households expected to have someone 15-35
- 2 – List stratum for households where the ages of respondents are unknown
- 3 – List stratum for households not expected to have someone 15-35

These strata should be properly identified within any statistical package for computing estimated variances from complex samples (SAS survey analysis procedures, SPSS with Complex

Samples module, STATA, etc.). Since data sets were constructed separately for each year, if study years are combined for analysis, the study year should also be indicated as the first level of stratification and the within year strata just described as the second level of stratification.

Two other factors need mentioning. The first involves the Year 1 data. As indicated, a design change from simple RDD to the stratified approach occurred during year 1. Years 2 and 3 simply always used the stratified approach. Because of the change during year 1, it may be more accurate to use a five stratum approach for this year. In this approach, the fifth stratum would be the RDD prior to stratification. This can be identified in the data as follows:

RDD data collected prior to stratification has the 0 code for stratum in the seventh digit of the id number AND has a first digit of 1 or 2.

RDD data collected after stratification has a 0 code for stratum in the seventh digit of the id number AND a first digit of 3.

The three list strata are identified as described earlier. This may be more accurate as it attempts to take into account the design change during year 1. Years 2 and 3 are unaffected by this.

The following table displays unweighted and weighted sample totals by stratum for each year. In the table the strata are defined as follows:

- 0 – RDD stratum used in stratification scheme after list samples have been removed
- 1 – List stratum of households expected to have someone 15-35 years old
- 2 – List stratum of households where ages of residents are unknown
- 3 – List stratum of households not expected to have anyone 15-35 years old
- 4 – RDD stratum in Year 1 before stratification was adopted. This stratum is determined as described above.

**Table 6: Unweighted and Weighted Sample Totals by Year and Stratum**

	Year 1		Year 2		Year 3	
Stratum	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
0	220	36,831	320	98,231	120	68,251
1	151	9,182	242	19,267	189	23,094
2	352	36,004	464	70,685	345	99,163
3	6	2,708	7	3,223	1	898
4	519	106,681	0	0	0	0
Total	1,248	191,406	1,033	191,406	655	191,406

The above table is useful for determining the unweighted sample sizes (completed interviews) by year and stratum and also for ascertaining an approximate average weight by year and stratum.

The final factor needing to be discussed regards the list stratum for households not expected to have anyone 15-35. This stratum was sampled sparsely for obvious reasons. Still, there were usually a few cases that showed up as eligible as the previous table displays. Within each year, there are 6 or fewer interviews and ever fewer cases that returned specimens within this stratum. In particular, in Year 3 there is a single interview and this person returned their specimen cup so there is a single returned cup. A stratum of size 1 can cause problems in variance estimation so it is recommended that this stratum be collapsed. Collapsing should only be done in Year 3. For this year, stratum 3 (the list stratum of households not expected to have anyone 15-35) should be collapsed with stratum 0 (the RDD stratum used in the stratification design). This should only be done for Year 3. Therefore, the net result is that there are 5 strata in Year 1, 4 strata in Year 2 and 3 strata in Year 3. As was stated earlier, it is always important to number the strata in such a way that both Year and stratum are known and accounted for within the statistical package being used.

## **VII. Survey Effort**

A great deal of effort was expended in attempting to get the highest response rates possible and the highest quality data possible. Table 7 summarizes this effort by examining the number of call attempts required to get various results.

**Table 7. Effort Expended to Complete Interviews**

Final Case Outcome	Avg. Number of Call Attempts	Maximum Number Of Call Attempts
No Contact – Unresolved Residential Status	13.6	35
Screening Limit	21.3	50
Screening Refusal	16.1	48
Household Not Eligible	3.7	48
Refusal After Eligible Respondent Identified	10.6	50
Limit After Eligible Respondent Identified	27.1	61
Completed Interview	6.4	50

Overall, cases were pursued until it was determined that additional efforts would not yield a more positive result.

### **VIII. Conclusions**

This was a very difficult study. In an era of declining response rates to telephone surveys, it was made all the more difficult. The study also combined the following elements that make a survey difficult to do:

- 1) Seeking younger age respondents (15-35 years of age).
- 2) Requiring parental permission (for 15-17 year olds).
- 3) Targeting a highly urbanized area (Baltimore city).
- 4) Targeting a highly minority population.
- 5) Having a very sensitive subject matter (sexual practices, drug use, etc.).
- 6) Requiring more than an interview (acceptance and return of a specimen).

Taken together, this study was one of the more difficult endeavors to attempt. However, as described throughout this report, the study was quite successful as denoted in the following results:

- 1) Successfully interviewing nearly 60% (58.7%) of all identified eligible respondents
- 2) Attaining an overall AAPOR Method 3 response rate of over 40% (40.4%)
- 3) Getting about 86% of all respondents to agree to accept a urine specimen cup
- 4) Getting over 92% of eligible ten respondents to agree to accept a specimen cup
- 5) Getting nearly 85% of all people who accept a cup to return it

- 6) Getting over 90% of parents to allow their teen to be interviewed
- 7) Getting over 84% of parents to allow their teen to be interviewed and offered the specimen cup

These rates suggest a protocol that produced high levels of cooperation amid a very difficult and challenging environment. They also suggest that a protocol such as this can be successful in other environments that may not be as challenging as the one described here.

## **IX. Acknowledgements**

This study would not have been possible without the cooperative efforts of a great many people at various institutions. Charles F. Turner of City University of New York (Queens College and Graduate Center), Susan M. Rogers and Elizabeth Eggleston of the Research Triangle Institute (RTI) managed the entire study and kept all pieces of it on track and under control. Their ability to keep everyone informed, handle a myriad of situations and keep the study always moving forward is greatly appreciated. Maria Villarroel and Sylvia Tan of RTI offered tremendous support in getting up to date status reports, creating data files of merged data from various sources and offering all kinds of analytic and computer support. Laxminarayana Ganapathi of RTI kept the TACASI computers running smoothly for three years with little interruption and was always available to troubleshoot problems and get quick solutions. Noelle Richa of RTI managed financial incentive records and James R. Chromy of RTI offered much needed statistical consultation. Marcia Hobbs of the University of North Carolina Chapel Hill supervised the lab testing of urine and swab specimens and Kimberly Rich and Alice Stillson communicated with everyone about the status of all specimen receipts and test results. At CSR, Ebony Haley supervised the clerical staff for mailing all letters and kits and did so keeping everything always current and with very few errors for the entire three years. Phyllis Doucette made sure we had all contractual matters and all supplies in order to carry out all the required study tasks. Rebecca Crow served as my assistant and kept all study processes under control including all weekly progress reports, keeping all clerical tasks and interviewing tasks running smoothly, programming the databases and statistical programs and generally making sure everything was smooth and timely. Without her dedicated and hard work, the study would simply not have been as successful. Susan Hynek managed all telephone interviewers and kept everyone motivated and working at a high level for three years. This is not an easy task. Finally, the CSR interviewing staff kept at it for three years always pushing to get better and better over time. Their efforts are truly appreciated. I personally wish to thank all of these people and to say it was a pleasure to work with such a talented team. As always, any errors in this final methodological report are mine alone.

## References

1. Roman, Anthony M.; Elizabeth Eggleston; Charles F. Turner; Susan M. Rogers; Rebecca Crow; Sylvia Tan; James R. Chromy; Laxminarayana Ganapathi. "Effects of Sampling and Screening Strategies in an RDD Survey", Proceedings of the 2008 Joint Statistical Meetings, Section on Survey Research Methods.
2. [HTTP:FACTFINDER.CENSUS.GOV](http://factfinder.census.gov).
3. Roman, Anthony M. "Survey of AIDS and Other Social Issues: Field Report", Submitted to the Research Triangle Institute, August 2000.