

**Supporting Statement B for Methodological  
Research to Support the NCVS: Self-Report Data on  
Rape and Sexual Assault: Feasibility and Pilot Test**

**February, 2014**

# Table of Contents

---

<u>Section</u>	<u>Page</u>
<b>B. Collection of Information Employing Statistical Methods.....</b>	<b>4</b>
B.1 Respondent Universe and Sampling Methods.....	4
B.1a..... Respondent Universe	4
B.1b..... Sampling Frames	6
B.1c..... Sample Design	13
B.2 Procedures for Information Collection.....	16
B.2a..... Rostering and Recruitment	16
B.3 Methods to Maximize Response.....	36
B.3a..... Address Based Sample	36
B.3b..... Telephone Procedures to Maximize Response	37
B.3c..... Calculation of Response Rates	38
B4. Test of Procedures or Methods.....	41
B4a..... Cognitive Testing	41
B4b..... Feasibility study	41
B5. Contacts for Statistical Aspects and Data Collection.....	42
<b>References:.....</b>	<b>44</b>

## **Table of Contents (Continued)**

---

<u>Tables</u>	<u>Page</u>
Table 1. Age-specific rates and standard errors for estimated rates of rape for females by age.....	5
Table 2. Approximate Respondent Universe for the Population Samples.....	6
Table 3. Selected characteristics related to RDD sample of the five metropolitan areas included in the study.....	10
Table 4. Number of exchanges, working banks, and telephone numbers in the landline and cell phone sampling frames by CBSA.....	13
Table 5. Sample Sizes and Assumptions for In-Person Interviews.....	14
Table 6. Assumptions and expected sample sizes for landline RDD sample by CBSA.....	17
Table 7. Assumptions and expected sample sizes for cell phone sample by CBSA.....	17
Table 8. Sample Design for the re-interview*.....	18
Table 9. Projected Rates and Standard errors for estimates of rape and sexual assault by data collection method and sample.....	27
Table 10. Approximate annual prevalence rates and standard errors for the NCVS for females 18-49 during 2005-2011.....	29
Table 13. Standard errors for the estimated unfounding rates, for true unfounding rates of .10 and .30.....	32
Table 14. Estimated Values and Precision of Kappa Statistics by sampled group.....	34
<u>Figures</u>	<u>Page</u>
Figure 1. Project Number of Completed Interviews by Sample Components for NSHS Pilot Survey	
Figure 2. Project Number of Completed Interviews by Sample Components for NSHS Feasibility Survey	
<u>Exhibits</u>	<u>Page</u>
Exhibit 1. Questionnaire content by treatment group	

## **B. Collection of Information Employing Statistical Methods**

The following section describes the statistical methods planned for the NSHS. In Section B.1, we describe the target population of the NSHS and the methods for sampling. In section B.2, we describe the procedures for collecting survey information; weighting and estimation procedures; an appraisal of the expected reliability of the estimates; and evaluation plans. Section B.3 describes methods to maximize response rates and the calculation of response rates. Section B.4 describes the tests of procedures and B.5 describes the contacts for design of the project.

### **B.1 Respondent Universe and Sampling Methods**

#### **B.1a Respondent Universe**

The goal of this study is to develop a methodology to collect estimates for rape and sexual assault (RSA) for the female civilian non-institutional population. To achieve this goal we will collect data on RSA using two different methods (in-person and telephone). The universe for the study is restricted to females living in one of five metropolitan areas, namely, the Core Based Statistical Areas (CBSAs) of Dallas-Fort Worth-Arlington, TX; Los Angeles-Long Beach-Santa Ana, CA; Miami-Fort Lauderdale-Pompano Beach, FL; New York-Northern New Jersey-Long Island, NY-NJ-PA; and Phoenix-Mesa-Glendale, AZ.

For cost efficiency, one of our two probability samples, the address-based sampling component, will target women age 18-49. Women over age 49 are at very low risk of RSA. Table 1 below shows the rates of RSA based on NCVS data.

Table 1. Age-specific rates and standard errors for estimated rates of rape for females by age

<b>Age group</b>	<b>Rate</b>	<b>Standard Error</b>
Age 12-17	0.0036 5	0.0005 3
Age 18-24	0.0031 1	0.0004 4
Age 25-29	0.0022 0	0.0003 9
Age 30-34	0.0015 9	0.0003 7
Age 35-39	0.0014 7	0.0003 9
Age 40-44	0.0008 8	0.0002 3
Age 45-49	0.0009 8	0.0002 4
Age 50-54	0.0005 0	0.0001 5
Age 55-59	0.0006 2	0.0003 0
Age 60-64	0.0003 6	0.0001 9
Age 65-74	0.0002 4	0.0001 1
Age 75-84	0.0000 4	0.0000 4
Age 85+	0.0003 3	0.0002 5
Age 18+	0.0011 8	0.0000 9

Source: Derived from analysis of the NCVS public use files for 2005-2010.

Considering this risk in conjunction with the relative cost of in-person screening and interviewing, targeting women age 18-49 provides the most cost-efficient way to maximize the measurement goals of the study. The telephone probability sample will target women age 18 and over. The telephone sample is not restricted to an age group because the relative cost of screening for a restricted age group by telephone does not make this an efficient way to meet the goals of the study.

Both samples target a geographic and demographic subset of the target population of the National Crime Victimization Survey (NCVS). Results from

both samples may be compared to the NCVS and to each other for women age 18-49 in these CBSAs, and the telephone component may be compared to the NCVS for ages 18 and over in these CBSAs. Table 2 provides Census 2010 population counts by age for females for each CBSA.

Table 2. Approximate Respondent Universe for the Population Samples

CBSA	Age Categories					
	18+ Years	18-29 Years	30-39 Years	40-49 Years	18-49 Years	50+ Years
Dallas	2,362,811	537,373	487,657	474,202	1,499,232	863,579
Los Angeles	4,966,994	1,144,799	921,548	944,716	3,011,063	1,955,931
Miami	2,281,225	427,339	373,801	428,179	1,229,319	1,051,906
New York	7,690,949	1,581,906	1,334,056	1,444,772	4,360,734	3,330,215
Phoenix	1,565,537	345,396	288,705	280,387	914,488	651,049

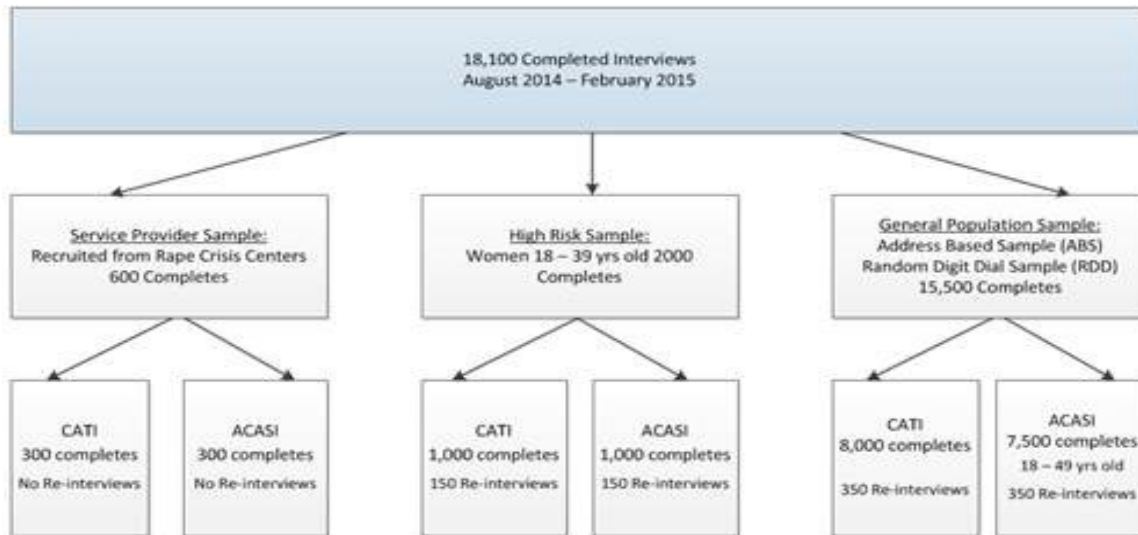
\*Based on results of the 2010 Census, SF1, Table P12, downloaded from the American Fact Finder on May 15, 2013, <http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml>. The counts include females in institutions, who are not part of the target universe. The counts of females age 18+ in institutions are 18,664, 29,937, 14,191, 73,562, and 10,823, in Dallas, Los Angeles, Miami, New York, and Phoenix, respectively, less than 1% in each case.

### B.1b Sampling Frames

The overall study comprises four components: 1) a probability sample using personal visit, 2) a probability sample using telephone, 3) a purposive sample of women age 18-39, and 4) a purposive sample of known victims. The majority of the NSHS resources will be invested in the two probability samples of the target population: females age 18 and over in the civilian, non-institutional population. One probability sample uses address based sampling to enable personal visit interviews, including ACASI. The other probability sample is for telephone interviewing and combines a list-assisted random digit dial (RDD) landline frame and a RDD frame of cell phone numbers. The in-person and telephone procedures will be compared to each other, as well as to the NCVS. The purposive sample of women 18-39 will be used to supplement the analysis comparing the in-person and telephone surveys. These women are at higher risk of assault and will increase the statistical power when conducting some analyses. The sample of known victims will be used to assess how the instrumentation works for individuals who are known to have been a victim of relatively serious assaults.

Figure 1 provides an overview of the different sample frames and the expected number of completed interviews from each frame. The targeted number of completed interviews is 18,100, with 1,000 re-interviews.

Figure 1. Project Number of Completed Interviews by Sample Components for NSHS Pilot Survey



**Address sample.** The address sample for the study will be selected using a three-stage probability sample design. Most of the sample will involve the selection of: (1) segments of dwelling units, (2) addresses in each sampled segment, and (3) a female respondent age 18-49 within the screened dwelling. A separate stratum will be formed of census blocks containing female students in college dormitories or a high proportion of women 18-29. In this stratum samples will be selected at twice the sampling rate. The frames to be used at each stage are described below.

The first-stage sampling units (referred to as segments) will be based on census-defined blocks. The frame of segments will be created based on the block data from the 2010 Census Summary File 1 (SF1), in the counties of the CBSA defined by OMB as of 2008. For each block, SF1 provides population counts by age and sex, household counts by type of household, group quarters population by sex by age by group quarters type, and housing unit counts by occupancy status. Two strata will be formed. One strata will identify blocks with a high concentration of young women if: 1) there are more than 20 female students in dormitories or 2) they have a population of 100 persons or more, of which at least 30% are females age 18-29. The treatment of these distinct blocks as a separate stratum will be discussed below. The second strata will be all other blocks.



This frame of blocks based on SF1 will be matched to block-level counts of dwelling units based on address-based sampling (ABS) frames derived from U.S. Postal Service (USPS) address lists. The USPS address lists, which are called Computerized Delivery Service Files (CDSFs), are derived from mailing addresses maintained and updated by the USPS. These files are available from commercial vendors. Recent studies suggest that the coverage of these lists is generally high for urban and large suburban areas, and sometimes reasonably high for parts of rural areas as well. The SF1 counts of housing units and the ABS counts should typically be similar in the five CBSAs, but substantial discrepancies at the block level may signal potential issues affecting the appropriateness of the measure of size. New construction since 2010 could explain ABS counts substantially larger than the census counts. Census counts considerably larger than the ABS counts could signal undercoverage of the ABS frame, caused by, for example, use of post office boxes as a replacement for direct mail delivery to the address. Because of potential inconsistencies between the census and ABS frame in geocoding dwellings to blocks, discrepancies at the individual block level will be assessed relative to neighboring blocks to detect whether a systematic pattern can be observed.

A measure of size (MOS) will be assigned to each block based on the SF1 count of women 18-49, adjusted proportionally upward if the ABS count indicates new construction since the 2011 census. Blocks with no population in 2010 will be included in the segment formation process to ensure that areas containing dwelling units (DUs) constructed after the 2010 Census are given an appropriate chance of selection. Within each CBSA, the block-level records will be sorted by county, tract, block group, and block number before creating the segments. A single block will be used as a segment if the MOS in the block exceeds 30. Neighboring blocks will be combined within a tract to reach either the required minimum MOS of 30 per segment or the end of the tract (segments will not cross tract boundaries).

In a few fringe counties of the CBSAs where the ABS frame appears particularly deficient, conventional listing will be employed. At the first stage, segments will be formed based only on the SF1 information.

At the second stage of selection, a sample of addresses will be selected within sampled segments. The sample of addresses will be selected from a combination of two sources: (1) the address-based sampling (ABS) frames derived from USPS CDSF address lists; and (2) in the few counties where the ABS frame is not used, address listings of the segments compiled by field data collectors.

The second source of addresses is created to take account of possible new construction and to improve coverage in general in ABS areas. A missed address procedure will be implemented for a subset of the sampled segments to check for addresses that are not represented by the ABS frame. Such missed addresses are those that should have been included on the original ABS listing but were not, or were constructed after the lists were compiled. When a segment is selected for this procedure, the entire segment will be canvassed by the data collector, and any newly identified addresses deemed as missed or new construction will be added to the address sampling frame, but only after a search of the ABS frame in neighboring segments confirms that the addresses are truly missed rather than simply incorrectly geocoded on the ABS frame. Depending on the selection probability of the segment to receive the coverage check, either all or a sample of the missed addresses identified within a sampled segment will be selected for the study.

At the third stage of selection, females aged 18-49 will be identified after rostering persons age 18 years and over in the household by age and sex. One woman will be selected at random from the eligible respondents, if any are in the household. See below for procedures on this rostering procedure.

For the stratum that identifies young women, blocks will be sampled at the same rate proportional to their MOS as other blocks. Females in the dormitories within the block as well as females in dwelling units, if any, will be sampled. The within block sampling rate will be doubled compared to blocks in the rest of the sample, so as to increase the expected sample size within these blocks. This approach recognizes the probable benefit from oversampling this presumed higher risk population, with at least two such

blocks sampled per CBSA. Interviewers will first reconcile the addresses provided by CDSF to determine if and whether the dormitory population is included. Non-dormitory addresses will be sampled in the usual way, but interviewers will be instructed to determine the most effective way to list the dormitory population, whether by unit designation (e.g., room number) or student name.

**Telephone sample.** The telephone sample for the pilot study will employ a dual frame approach, consisting of a landline telephone sample and a sample of cell phones. The incorporation of a cell phone sample with the traditional landline sample is necessary to ensure substantially unbiased coverage of all telephone households. The populations for the five CBSA's are provided in Table 3. These sampling frames are described below.

**Table 3. Selected characteristics related to RDD sample of the five metropolitan areas included in the study**

Characteristic*	Metropolitan area [1]				
	Dallas TX CBSA	Los Angeles CA CBSA	Miami FL CBSA	New York NY CBSA	Phoenix AZ CBSA
Total number of households [2]	2,127,000	4,152,000	2,007,000	6,751,000	1,447,000
Percent of Households with telephones [3]	98.1%	98.4%	98.1%	97.1%	96.9%
Landline only	9.4%	12.8%	11.9%	21.6%	8.9%
Landline and cell phone	45.5%	68.6%	59.1%	56.4%	57.7%
Cell phone only	43.2%	17.0%	27.1%	19.1%	30.3%
None	1.9%	1.6%	1.9%	2.9%	3.1%
Percent of households with one or more females 18 or older [4]	84.2%	84.7%	84.1%	85.2%	83.6%
Total number of females 18 or older	2,325,700	4,814,700	2,223,300	7,626,400	1,520,900

[1] Metropolitan areas are Core Based Statistical Areas (CBSAs) defined by OMB.

[2] Source: 2005-2009 American Community Survey Five-Year Estimates.

[3] Source: Blumberg, et al (2011).

[4] Source: Derived from tabulations of the 2005-2009 American Community Survey Five-Year Estimates PUMS.

The method of random digit dialing (RDD) referred to as “list assisted RDD sampling” (Tucker, Lepkowski, & Piekarski, 2002; Brick and colleagues, 1995) will be used to sample households with landline telephones. To supplement the traditional list-assisted RDD sample, a second sample will be drawn from a frame of cell phone numbers. The “overlap” method (Brick et al, 2011), where a household is included regardless of whether it also has a landline, will be used to sample cell phone numbers. The overlap method: (1) accounts for the non-coverage of household with only cellular service, (2) is less costly than screening for households that use cell phones only (the “screener” method), (3) is a good way to reach households that have a landline but use their cell phone for most calls. The proposed dual frame approach requires appropriate procedures to weight the survey data (see section B.2c). We are targeting approximately 50% of the completed interviews will be done by cell phone and 50% will be by landline.

For the list-assisted landline RDD sample, special procedures will be implemented to reduce costs and to increase the efficiency of the sample.

The method starts out by specifying all possible 100-banks of telephone numbers that cover the particular geographic area of interest, where a 100-bank is defined to be the set of all telephone numbers with the same first eight digits (area code, exchange, and first two digits). From this set of telephone numbers, a stratified random sample is selected. We will then utilize a procedure that will reduce the number of unproductive calls in the landline RDD sample by removing any nonresidential business telephone numbers found in the White and Yellow Pages of telephone directories. These telephone numbers will be coded as ineligible and will not be released for telephone interviewing. In addition, prior to dialing the sample we will use an automated procedure in conjunction with manual calling to identify nonworking numbers. All sampled landline RDD telephone numbers, including those listed in the White Pages, will be included in this procedure. The numbers that are identified as nonworking numbers will also be coded as ineligible and not released for telephone interviewing. This procedure will also identify cell phone numbers that have been ported from landline exchanges. These telephone numbers will be treated as ineligible for the landline sample; however, they will be retained in the sample for use as a part of the cell phone sample described below.

Similar to the selection of landline telephone numbers, a sample of cell phone numbers will be selected from a frame consisting of blocks of numbers assigned to cell phones. For the cell phone sample, it is not possible to prescreen the sampled numbers using the procedures available for landline telephones. This is because directories do not exist for cell phones and predictive dialing of cell phone numbers is prohibited by law. Thus, unlike the landline samples where a substantially reduced set of potentially eligible telephone numbers is dialed, the full cell phone sample is typically dialed. However, data are available in the sampling frames that could potentially improve the efficiency of cell phone samples. The first is an indicator of the activity status of the cell phone number. Preliminary analysis of these data indicates that the vast majority of “no activity” cases are nonresidential telephone numbers while those cases without a zip code are also more likely to be nonresidential. These indicators may be beneficial in creating sampling strata within the cell phone sample. If it appears that the efficiency of the sample could be increased by oversampling those cases

that are more likely to yield a residential phone number, consideration will be given to employing these indicators to stratify the cell phone sample.

To balance the need to achieve the best possible coverage of telephone numbers in the CBSAs against improved sampling efficiency obtained by reduced screening, certain exchanges will be excluded from the sampling frame. For the landline RDD sample, exchanges with the lowest number of in-area households will be excluded. For the cell phone sample, excluding certain exchanges is more complicated as exchange assignment is more likely an indicator of where the phone was purchased not where its user resides. A recent development in cell phone sampling, that can be used to better identify exchanges that are more likely to be in the selected CBSAs, is the use of switch locations<sup>1</sup>. Using switch locations that are within the CBSAs boundaries will allow us to target users that are most likely to be in the geographic area of interest. Since the use of these methods is rapidly changing, we will be assessing the utility of this method prior to drawing the sample. For purposes of planning at this point, we have assumed that we will not use these methods.

For each of the 5 CBSAs, the estimated sampling frame size, coverage rate and out-of-area rate for the landline RDD sample is shown in Table 4. It can be seen that the coverage rates (after excluding those exchanges with small numbers of in-area households) is virtually 100 percent for all five areas. This table also summarizes the numbers of 1,000-cell phone blocks and corresponding counts of telephone numbers in the cell phone frame for these CBSAs.

**High Risk Group.** Averaged over the last 5 years, the NCVS data indicates a substantial age differential in reported incidents of rape, with the highest levels for women age 18 and over occurring in the 18-39 age range and declining with age thereafter (see Table 1a).<sup>2</sup> The third component of the RSA adopts an experimental approach to studying differences between the

---

1 <http://www.m-s-g.com/CMS/ServerGallery/MSGWebNew/Documents/GENESYS/whitepapers/Cellular-RDD-Frame-Construction.pdf>

2 Note the general population sample recruited for in-person contact is restricted to ages 18-49. The age group for the High Risk group was lowered to a maximum age of 39 to focus more sharply on the women who are at highest risk of rape.

two interviewing methods by targeting young women in the 18-39 age range. Young women will be recruited for the study as volunteers, with the requirement that they agree in advance to participate through either the personal visit or telephone modes by supplying contact information for both. The volunteers will be randomly assigned to mode, enabling a strict experimental comparison of the performance of the questionnaires and measurement protocols for this group of women, who are statistically at higher risk of recent victimization. The recruitment methods do not focus specifically on rape or sexual assault, so the majority of these respondents are not expected to report victimizations. The methods for recruitment of this group are described later in this document.

**Service Provider Sample.** A fourth component, the Service Provider sample, will be based on recruiting women recently seeking support from rape crisis centers. This component again allows an experimental comparison between the modes. Many of these respondents will have incidents to report, and the emphasis in this component of the NSHS will be to obtain qualitative data on the flow of the questionnaires and any difficulties the respondents may encounter.

**Table 4. Number of exchanges, working banks, and telephone numbers in the landline and cell phone sampling frames by CBSA**

RDD Sampling Frame	Metropolitan area (CBSA)				
	Dallas TX CBSA	Los Angeles CA CBSA	Miami FL CBSA	New York NY CBSA	Phoenix AZ CBSA
Landline Frame					
Total number of exchanges	1,443	2,748	1,213	4,418	778
Exchanges included in sampling frame	1,406	2,702	1,197	4,330	706
Number of working 100-banks included in frame [1]	56,448	109,943	59,637	196,389	35,073
Total number of telephone numbers in frame	5,644,800	10,994,300	5,963,700	19,638,900	3,507,300
Coverage of sampling frame [2]	99.95%	99.93%	99.99%	99.99%	99.90%
Out-of-area rate [3]	0.36%	0.34%	0.19%	1.07%	0.70%
Cell phone frame					
Number of 1,000-series banks	8,584	17,050	8,018	27,158	5,410
Total number of telephone numbers in frame	8,584,000	17,050,000	8,018,000	27,158,000	5,410,000

[1] Working banks with at least 1 listed telephone number.

[2] Estimated percentage of landline households in the exchanges included in the sampling frame.

[3] Percent of landline households in frame that are located outside of metropolitan area.

### B.1c Sample Design

**Address sample.** As described earlier, the sample will be selected using a three-stage, stratified probability design which will result in a total sample of about 7,500 female respondents aged 18-49, equally divided among the five CBSAs in the study. The sample will include both persons living in households and in civilian, non-institutional group quarters.

Within each CBSA, a sample of segments consisting of blocks or groups of blocks will be drawn. To control for varying concentrations of the target population, the segments will be stratified into four strata in each CBSA by the ratio of the MOS to the number of dwellings in the ABS frame. Within each stratum, the frame will be sorted by county and tract, and a systematic PPS sample of segments will then be selected from the sorted frame, with probability proportional to the MOS.



At the second stage of sampling, current lists of addresses will be developed for the selected segments using address lists compiled from the CDSF lists. From these lists, samples of addresses will be drawn for screening using systematic sampling within the segments. As a result of the PPS selection of segments, the number of addresses selected for screening will vary somewhat, but the expected number of eligible female respondents should be approximately equal, thus maintaining an approximately constant workload across segments. A screener interview will be administered within the household associated with each sampled address to determine the sex and age of each household member, and the active duty status of military personnel.

The third and final stage of selection will use the roster of age-eligible household members (females aged 18-49) for each dwelling unit associated with a selected address (or household-equivalent in the case of group quarters) to select a single female respondent with equal probability. The sampling rate will be adjusted for the average number of females 18-49 in households where any are present, to avoid a shortfall in the sample from this aspect of the sampling. The ACS data for 2006-2010 gives the average number of women age 18-49 in households with any such women as 1.12, 1.20, 1.15, 1.17, and 1.13 in the Dallas, Los Angeles, Miami, New York, and Phoenix CBSAs, respectively.

Table 5 provides the anticipated sample to be drawn, the eligibility and response rates assumed for the final design. The eligibility rates are taken from census data. Response rates are based on field studies that Westat is currently conducting for other federal clients. These assumptions include a 70% response rate to complete the household screening interview and 70% response from women who are selected to participate in the survey. It is anticipated that about 52% of the screened households will have an eligible female.

Table 5. Sample Sizes and Assumptions for In-Person Interviews

ACASI interviews	7,500
ACASI response rate	70%
Completed HH screeners (with eligible respondents)	10,714

Occupancy rate	89%
Eligibility rate	52%
Screener response rate	70%
DUs sampled	33,073

**Telephone sample.** Once the landline and cell phone sampling frames have been specified, simple random or systematic samples of telephone numbers will be selected from each of the five CBSAs. To ensure that telephone numbers in certain types of exchanges are appropriately represented in the sample, the telephone numbers in the landline sampling frame will be stratified (either explicitly or implicitly through sorting) by relevant exchange-level characteristics prior to selection. The cell sampling frame does not have exchange based characteristics, so exchange-based stratification of the cell sample is not possible. However, it will be possible to sort these by county.

The exchange-level data that will be used for stratification in the landline sampling frame are derived from Census population counts and include percentages for the following characteristics:

- Age - 0-17; 18-24; 25-34; 35-44; 45-54; 55-64; 65+
- Race - White; Black; Hispanic; Asian/Pacific Islander
- Income - 0K-<10K; 10K-<15K; 15K-<25K; 25K-<35K; 35K-<50K; 50K-<75K; 75K-<100K; 100K+
- Home - Owners; Renters/Other
- Education - College Graduates

Tables 6 and 7 show the expected levels of screening needed to achieve 800 interviews per CBSA from each of the landline RDD and cell phone frames. The assumptions provided in these tables are based on our past experiences. For the landline sample we utilize data specific to each CBSA, while for the cell phone sample our assumptions are based on our overall experience with cell phones and are not specific to the CBSAs.

**Re-interview sample.** We are proposing to re-interview a sample of respondents who report being victimized in their lifetime. This will focus the analysis on the consistency of initial reports of victimization at both the screener and the detailed incident form. The re-interview will repeat the same protocol as initially completed, using the same reference period. We are proposing the re-interviews be scheduled approximately 2 weeks after the original interview. This leaves enough time for the memory of the interview to fade, but does not wait too long to confound reference periods.

In addition to excluding those who do not report a victimization, the re-interview will also exclude:

- Respondents in the Service Provider sample
- CATI respondents who are age 50 years old and over. This is to maintain comparability between the CATI and ACASI arms
- Anyone who exhibits extreme distress during or after the interview

- Those who have had a past 12 month incident but are unwilling to answer detailed incident items, or who break off the interview while answering those items.

The goal is to complete a total of 1,000 re-interviews, including 500 re-interviews in each mode. Among these 1,000, we will include:

- All women reporting victimization by rape or sexual assault within the last 12 months and who either fill out at least one detailed incident form or say they cannot recall the details of the incident,
- A sample of women reporting lifetime exposure a year or more prior to the interview.

Of the 500 completes in each mode, 350 are allocated to the general population sample and 150 are allocated to the high risk sample.

For design purposes, we have worked with recent rates from the British Crime Survey (BCS) for the general population sample (Hall and Smith, 2011). The high risk sample was assumed to have twice the risk during the preceding year as the population sample. We have assumed that 80% of the respondents to the initial phase of NSHS would participate in the re-interview if sampled. Using the rates from the BCS, Table 8 provides the proposed sample design.

## **B.2 Procedures for Information Collection**

### **B.2a Rostering and Recruitment**

#### *General Population Field (ABS) Sample*

For the ABS sample, we will initially send a mail survey to each sampled address. The survey will ask the respondent to provide a roster of adults who live in the household. Attachment B provides the letter that will be included in the package and Attachment C provides the mail survey. A reminder postcard will be sent to all households one week later, and a second, replacement roster will be sent to households that do not return the first roster (Attachments B-1 and B-2).

The advance letter is not specific about the purpose of the study in order to keep the content of the survey confidential. This protects a female in the house from a perpetrator who might be living with her. It is not until the

selected respondent agrees to participate and the respondent is in a private location that more information is provided about the survey. This procedure is identical to what has been used on prior, general population, studies that have been conducted.

Table 6. Assumptions and expected sample sizes for landline RDD sample by CBSA

Sampling unit	NYC CBSA		Dallas CBSA		LA CBSA		Phoenix CBSA		Miami CBSA	
	Rate [1]	Exp. no.	Rate [1]	Exp. no.	Rate [1]	Exp. no.	Rate [1]	Exp. no.	Rate [1]	Exp. no.
Telephone numbers	---	28,516	---	20,108	---	17,540	---	21,046	---	38,766
Nonresidential/nonworking [2]	54.0%	15,408	57.2%	11,501	49.9%	8,757	59.3%	12,483	54.0%	20,929
Numbers available for telephone screening		13,108		8,607		8,783		8,563		17,836
Nonresidential/nonworking	24.2%	3,177	29.8%	2,566	21.3%	1,875	20.1%	1,720	33.9%	6,046
Residency status undetermined	40.3%	5,283	33.7%	2,899	37.9%	3,331	39.9%	3,419	39.0%	6,958
Determined to be residential	35.5%	4,649	36.5%	3,142	40.7%	3,577	40.0%	3,424	27.1%	4,832
Households completing screener	28.8%	1,340	43.2%	1,357	37.7%	1,349	39.9%	1,367	28.1%	1,359
Households with adult female	85.3%	1,143	84.2%	1,143	84.7%	1,143	83.6%	1,143	84.1%	1,143
Sampled female	1 per HH	1,143	1 per HH	1,143	1 per HH	1,143	1 per HH	1,143	1 per HH	1,143
Respondent (completed interview)	70.0%	800	70.0%	800	70.0%	800	70.0%	800	70.0%	800

[1] Assumptions based on experience in prior RDD surveys conducted by Westat.

[2] Identified through pre-screening processes prior to calling by the TRC.

Table 7. Assumptions and expected sample sizes for cell phone sample by CBSA

Sampling unit	NYC CBSA		Dallas CBSA		LA CBSA		Phoenix CBSA		Miami CBSA	
	Rate [1]	Exp. no.	Rate [1]	Exp. no.	Rate [1]	Exp. no.	Rate [1]	Exp. no.	Rate [1]	Exp. no.
Telephone numbers	---	22,247	---	22,247	---	22,247	---	22,247	---	22,247

Nonresidential/nonworking [2]	38.1%	8,480	38.1%	8,480	38.1%	8,480	38.1%	8,480	38.1%	8,480
Residency status undetermined [2]	19.1%	4,244	19.1%	4,244	19.1%	4,244	19.1%	4,244	19.1%	4,244
Determined to be residential [2]	42.8%	9,524	42.8%	9,524	42.8%	9,524	42.8%	9,524	42.8%	9,524
Eligibility Rate	80.0%	7,619	80.0%	7,619	80.0%	7,619	80.0%	7,619	80.0%	7,619
Cell phone belongs to a female	50.0%	3,810	50.0%	3,810	50.0%	3,810	50.0%	3,810	50.0%	3,810
Females completing screener	30.0%	1,143	30.0%	1,143	30.0%	1,143	30.0%	1,143	30.0%	1,143
Respondent (completed interview)	70%	800	70.0%	800	70.0%	800	70.0%	800	70.0%	800

[1] Assumptions based on experience in prior RDD surveys conducted by Westat.

[2] Identified through telephone screening.



Table 8. Sample Design for the re-interview\*

	Report in the Last 12 months Sampled+	Report Lifetime but not within last 12 months		
		Total **	To Be Sampled for Re-interview	Reciprocal sampling rate
ACASI, general	234	1248	116	10.76
ACASI, high risk	62	135	88	1.54
CATI, general	153	817	197	4.15
CATI, high risk	62	135	88	1.54

\* Prevalence rates are based on results from the British Crime Survey (Hall and Smith, 2011). For simplicity in illustrating the effect of the re-interview, all figures in the table have been adjusted for an 80% response rate to the re-interview. For example, the BCS results suggest that the ACASI sample may yield 292 persons reporting rape or sexual assault during the previous year, but only 234 are projected to participate in the re-interview.

+ All persons reporting an incident within last 12 months who complete a detailed incident form or are willing to complete a detailed incident form but do not recall enough details to answer the questions will be included in the re-interview sample with certainty.

\*\* Expected number from the main survey.

Households responding to the mail roster that do not have an 18 to 49 year old female will be coded as ineligible and no interviewer visit will be conducted. All other households (i.e., those not responding and those with an eligible female) will be visited by an interviewer. If the household did return a roster and it was determined that an eligible adult female age 18-49 resides there, a letter will be mailed to the household informing them that their household is eligible for the study and that an interviewer will be visiting soon (Attachment D). Respondent selection will be done at the time of the visit.

For households that do not complete the mail roster, the interviewer will conduct the household rostering at the time of the initial visit, which requests age and gender of adults in the household in order to randomly select an adult female ages 18-49. (Attachment S).

### *General Population Telephone*

Procedures for the screening of telephone sample are as follows:

Landline: For the landline sample, we will conduct a reverse address match on the sample of phone numbers. If an address can be identified, an advance letter will be sent to the household explaining the purpose of the study, and informing the household that an interviewer will be calling in the coming days. (Attachment E)

The interviewer will confirm that the household is located within the CBSA by asking for the county and will then screen the household to randomly select an eligible female adult. (Attachment T)

Cell Phone Sample: For the cell phone sample, we will assume that the cell phone is not shared with others. Thus, if a male answers the cell phone, the interviewer will thank the respondent and terminate the call. If a female answers, the interviewer will confirm that she lives within the CBSA and is aged 18 or over. (Attachment T)

Households for which we have an address and that refuse in a non-hostile manner will be mailed a refusal conversion letter asking that they reconsider their willingness to participate (Attachment F).

### *Recruiting the High Risk Sample*

To recruit the high risk sample, first we will seek cooperation of Institutions of Higher Learning (IHL) in each of the CBSAs (Attachment G) to allow us to post recruitment materials on their campuses or electronic boards as well as through online sources such as Craigslist.. Women 18-39 will be asked to participate. Those 30-39 will be sampled as they are needed to fill in the quota for the sample. All recruitment material will display options for respondents to indicate interest in study participation by emailing, by web, or by calling a toll-free number. When calling, they will be asked to leave a voice mail. Messages will be monitored daily by staff (Attachments H, I). Eligibility screenings will be self-administered for those responding by the web, or administered by an interviewer for respondents that chose to call the provided phone number or email with their contact information (Attachment J).

If a respondent is found to qualify for the study, they will be randomly assigned to either the field or telephone design condition. An interviewer will arrange a time to conduct the interview. The in-person interviews will primarily be conducted in the respondent's home unless the respondent requests to meet elsewhere. The script to actually schedule a time for the interview is provided in Attachment K.

### *Recruiting Service Provider Sample*

Both the Feasibility and Pilot Studies will include a Service Provider sample of participants known to be survivors of sexual assault from each of the five CBSAs.

Agencies that have contact with survivors will be recruited to assist with the study, primarily rape crisis centers (RCC). (Attachment L) These agencies will be asked to share information about the study with survivors and to provide space for interviews, if feasible.

Materials used to recruit this sample will depend on the preferences of the RCC, but may include flyers or direct communication (email, mailings, etc.) between RCC employees and clients. Volunteers will be asked to call an 800 number to get more information about the study and to set an appointment. The flyer that will be distributed and the script to make appointments with this group are provided in Attachment M.

Field interviewers will work carefully to find a safe and convenient location for the interview. The first choice will be the service provider from where the individual was recruited. Otherwise it could be done in a meeting room at a local public library or other similar space that allows for both safety and privacy.

Telephone respondents will be contacted by the interviewer, who will first confirm that the respondent is in a private location where no one else might be listening to the interview.

## **B.2b Data Collection**

The study will interview approximately 18,100 respondents. A subset of 1,000 will be contacted approximately two weeks later to do a re-interview. The re-interview will be identical to the first interview (Figure 1).

The CAPI portion of the in-person visits and all of the CATI interviews will be audio-recorded when there are no objections from the respondent.

Once the respondent has been selected, she will be informed of her rights as a human subject. In-person respondents will be asked to read a consent form privately on the laptop. The respondent will click on the computer screen to acknowledge consent, and the interviewer will enter an ID to acknowledge consent (Attachment S). Telephone respondents will be provided the consent orally (Attachment T).

An overview of questionnaire content is provided in Exhibit 1. The full in-person questionnaire is provided in Attachment S, and the CATI questionnaire is provided in Attachment T. Frequently asked questions for householders (prior to respondent selection) and for the selected respondent are provided in Attachment P and Attachment Q.

Exhibit 1. Questionnaire content by treatment group

<b>Survey content</b>	<b>ACASI</b>	<b>CATI</b>
Demographics	YES	YES
Event History Calendar	YES	NO
ACASI Tutorial	YES	NO
Victimization Screener	YES	YES
Detailed Incident Form	YES	YES
Vignettes	YES	YES
Debriefing	YES	YES
Distress Check	YES	YES
Provision of resources	YES	YES
Incentive payment	YES	YES
Re-interview request (if selected)	YES	YES

Apart from the event history calendar, which is presented via paper to in-person respondents in order to assist them in dating events (Attachment O),

and the tutorial for the ACASI system (Attachment S), the content of the interview will be the same across modes. The order of the questions in the Victimization Screener differs by mode. The content and order of the questions in the Detailed Incident Form (DIF) are identical. The primary difference in the DIF is the CATI questions are structured to primarily ask for responses that are either yes/no or ask for a number. This is to protect the confidentiality of the respondent in case someone in the household is within earshot of the respondent.

The interview begins with a short series of demographic and personal items (Attachments S and T). A series of victimization screening questions is then administered to determine if the respondent has experienced 12 types of rape and sexual assault in their lifetime or in the past 12 months. Following prior studies, the screener cues respondents with explicit reference to behaviors that make up the legal definition of rape and sexual assault. The telephone protocol screens first for lifetime experiences, then for whether a reported incident occurred most recently in the past 12 months. ACASI respondents are asked first if they have had an experience in the past 12 months, and if not, are later asked if they have ever had such an experience in their lifetime.

Those with one or more incidents in the past 12 months continue into the Detailed Incident Form, which asks for more information about the circumstances leading up to, during, and following the incident(s). Based on past results, we expect that about 5% of women will report some type of incident occurring in the last 12 months. About 1% or less (depending on the estimate) will report a rape or attempted rape (i.e., something involving penetration). The instrument will be programmed to cap the number of detailed incident forms to three incidents in the past 12 months. It gives priority to the incidents involving rape.

Following the detailed incident form (or following the screener if no incident is reported), the respondent will be presented with two vignettes which characterize different levels of coercion or alcohol use and are asked to answer survey questions about the vignettes. After the vignettes, the respondent will be asked a short series of 10 debriefing items to assess their

experience in completing the survey. These questions address distress and opinions about the survey using modified items from the Reactions to Research Participation Questionnaire (RRPQ), items to detect any portions of the questionnaire that were difficult to understand, and in the re-interview, a short set of items to determine the utility of the resources provided at the end of the first interview.

At the conclusion of the interview, (after the field interviewer has collected the laptop from the respondent) the field and telephone interviewers will check the distress level of the respondent (see section C below), and assuming it is safe to conclude the interview, will offer local resources to the respondent, pay the respondent (or collect her mailing address to mail the check), and if selected for re-interview, set up a time for the re-interview.

### *Interviewer Selection and Training*

We will prioritize the selection of interviewers who have administered similarly sensitive surveys previously (e.g., surveys with cancer patients). Phone interviewers may also be hired who have specific experience with the topic of sexual assault. All interviewers, regardless of location and background experience, will receive the same training on study procedures, survivor populations, and protocols for dealing with participant distress.

### *Quality Control*

During the data collection period, numerous quality control procedures will be utilized to ensure that data collectors are following the specified procedures and protocols and that the data collected are of the utmost quality. Approximately 10% of the telephone interviews will be monitored on a real time basis. Computer Assisted Recorded Interviews (CARI) will be used to monitor and validate interviewers in the field. Throughout the field period, supervisors will remain in close contact with the data collectors. Scheduled weekly telephone conferences will be held in which all non-finalized cases in

the data collector's assignment will be reviewed, to determine the best approach for working the case and the need for additional resources.

Management staff at all levels will have access to a supervisor management system, including automated management and production reports that will be used to monitor the data collection effort and ensure that the data collection and quality control goals are being attained. Data collectors will be required to transmit data on a daily basis. Data will be transmitted to a secure server at Westat's Rockville offices, which will then be used to update the automated management reports. These data are also used to produce weekly reports that might provide evidence of suspicious data collector behavior, such as overall interview administration length, individual instrument administration time, amount of time between interviews, interviews conducted very early in the morning or late in the evening, and number of interviews conducted per day.

### **B.2c Weighting and Nonresponse Adjustment**

**Address Sample.** The estimates from the NSHS will be functions of weighted responses for each sampled person, where the weights will consist of seven components: a DU base weight, three multiplicative adjustment factors (a household screener non-response adjustment factor, a within household selected person factor, and a selected person non-response adjustment factor) and a final post-stratification ratio adjustment. The weights for group quarters, particularly for college dormitories, will be derived in a generally similar manner. The details are as follows:

1. The DU base weight

The dwelling unit base weight is the inverse of the overall probability of selection of a sample housing unit or person in a group quarters. The probability of selection is the product of the conditional probabilities of selection at each stage of sample selection, where the stages are as follows: 1) the selection of segments in each CBSA; and 2) the selection of addresses within segments.

DUs sampled as part of the missed address procedure will be assigned similar base weights depending on the manner in which they were selected. A sample of segments will be selected for listing, and the DUs listed that were not part of the original ABS frame in the tract or neighboring blocks will be either all included in the sample or subsampled for inclusion. The base weight in this case must reflect the initial probability of the segment, the sampling for listing, and any subsampling of the missed addresses.

2. The household screener nonresponse adjustment factor  
The household screener nonresponse adjustment will be calculated to account for inability to obtain a completed household roster resulting from a refusal, failure to identify a knowledgeable screener respondent, or inability to locate the housing unit. Adjustment cells will be defined by grouping segments together by CBSA status, region, minority status, and other characteristics. For each cell, the ratio of the weighted number of eligible sample households to the weighted number of completed screeners will be computed, where the weight used will be the DU base weight multiplied by the first two multiplicative adjustment factors. This factor will be used to inflate the weights of the screener respondents in the cell to account for the screener nonrespondents. This adjustment is the first of the multiplicative factors that will be applied to the DU base weight for all completed screeners.
3. The within-household selected person weight  
One eligible person per household will be selected at random in households where the screener detects more than one eligible respondent. The multiplicative factor is simply the number of eligible respondents in the household. This is the second of the multiplicative factors that will be applied to the DU base weight.
4. The selected person non-response adjustment factor  
The selected person non-response adjustment factor accounts for those persons who were selected for the study, but for whom no ACASI interview was obtained. This type of non-interview can result from a refusal by the sample person, the inability to contact the sample person, etc. The factor will be computed using adjustment cells defined by relevant segment-level and screener data, where the weights of interviewed persons in the cells are inflated to account for the non-interviewed persons. This is the third of the multiplicative factors that will be applied to the DU base weight.
5. Post-stratification ratio adjustment  
This adjustment is designed to ensure that weighted sample counts agree with independent estimates of the number of women 18-49



in the civilian, non-institutional population of each CBSA for broad age groups. The adjustment may be expected to reduce both the bias and the variance of the estimates.

**Telephone Sample.** The weights for NSHS will be the product of a series of sequential adjustments. The starting point, for landline RDD and cell phone samples, will be a base weight that is defined as the inverse of the probability of selecting a telephone number from the sampling frame. In the early stages of weighting the landline and cell phone samples will be weighted separately by applying the appropriate weighting adjustments.

To create the person weight for the landline RDD sample, first a household weight will be created by adjusting the (telephone) base weight to account for unknown residential status, screener nonresponse, and multiple telephone numbers in household. The resulting household level weight will then be adjusted further to create a person level weight. The following factors are used to create the person level weight: the selection probability of the person within the household and the extended interview response rate. Similar procedures will be applied to calculate the person weight for the cell phone sample except that we will not have a person level adjustment as there will be no subsampling of eligible persons in the cell sample.

Since the landline and cell phone populations and samples overlap and the both samples are probability samples, we will use a multiple-frame estimation approach to account for the joint probabilities of selection from the overlapping frames. This approach follows the ideas of Hartley (1962).

There are three population domains of interest that we will account for when adjusting for the overlapping frames: (1) women in households with only landline service, (2) women in cell phone only households, and (3) women in households with both landline and cell phones. The landline RDD sample produces unbiased estimates for women in the first domain, while the cell phone sample produces unbiased estimates for women in the second domain. For women in the third (overlapping) domain a composite weighting factor  $\lambda$  ( $0 \leq \lambda \leq 1$ ) will be attached to the weights of women selected from the landline frame who also have a cell phone. Women from the cell frame

that have a cell phone will have a composite weighting factor of  $1-\lambda$ . The value of  $\lambda$  will be chosen to minimize the bias of the estimates for this domain (Brick et al., 2011).

After the landline and cell samples have been combined and person-level weights have been computed, we will perform post-stratification (Holt and Smith, 1979) to adjust the weights to known population counts. This procedure uses data from external sources (control totals) such as the American Community Survey (ACS) and its objective is to dampen potential biases arising from a combination of response errors, sampling frame undercoverage, and nonresponse. Here we use the term post-stratification loosely and intend it to include raking, a form of iterative multidimensional post-stratification (see Brackstone and Rao, 1979). In NSHS, the control totals will be derived primarily from demographic and socio-demographic data reported in the American Community Survey.

## **B.2d Standard Errors and Confidence Intervals for General Population Estimates**

Upon completion of data collection and processing, we will estimate the standard error,  $SE(\hat{p})$ , of an estimated prevalence estimate,  $\hat{p}$ , of an assault rate with replication methods that will account for the complex sample designs of the telephone and in-person surveys. We will use the estimated standard errors to produce confidence intervals of the form  $\hat{p} \pm t_{1-\alpha/2}(df) \times SE(\hat{p})$ , where  $t_{1-\alpha/2}(df)$  is the appropriate percentile of the t distribution on  $df$  degrees of freedom. The degrees of freedom,  $df$ , will depend on the choice of the replication method. The replication methods will support estimation of standard errors for other estimates as well, such as for the difference in estimated prevalence between the two surveys, different types of assault, and estimates for subdomains.

For purposes of planning the study, we use the approximation

$SE(\hat{p}) = \sqrt{\frac{DEFF p(1-p)}{n}}$ , where  $n$  is the unweighted sample size and  $DEFF$  is

the design effect reflecting the effect of the complex sample design. It is possible to approximate *DEFF* in advance by combining information about the intended design with experience from past surveys of a similar nature. The address sample is designed to collect data from an average of 5 respondents per segment. An intra-class correlation of .05, which has been previously observed for within-PSU violent crime in the NCVS, would suggest a design effect (DEFF) of 1.20. Variation in the weights due to non-interviews adjustments and within-household subsampling could raise this by an additional factor of 1.2, giving an overall DEFF of 1.44. Similarly, a design effect of 1.4 has been empirically observed from past telephone studies.

Table 9 provides estimated standard errors for the estimates from the two samples for two different outcomes, Rape and Other Sexual Assault. The general prevalence rates for these estimates were based on previous studies (Tjaden and Thoennes, 1998; Hall and Smith, 2011.) The overall results for the telephone survey are shown for the whole sample as well as for the expected proportion age 18-49, which will be used in the comparison to the in-person interviews.

Table 9. Projected Rates and Standard errors for estimates of rape and sexual assault by data collection method and sample

	In-Person		Telephone		Telephone 18-49	
	General Population	General Population plus high risk	General Population	General Population plus high risk	General Population	General Population plus high risk
Rape						
Rate	0.00450	0.00563	0.00300	0.00367	0.00450	0.00563
Standard Error	0.00093	0.00113	0.00072	0.00075	0.00113	0.00124
% CV	20.6	20.2	24.1	20.6	25.2	22.1
Other Sexual Assault						
Rate	0.03150	0.03938	0.02100	0.02567	0.03150	0.03938
Standard Error	0.00242	0.00293	0.00190	0.00197	0.00296	0.00322
% CV	7.7	7.5	9.0	7.7	9.4	8.2
Total						
Rate	0.03600	0.04500	0.02400	0.02933	0.03600	0.04500
Standard Error	0.00258	0.00312	0.00202	0.00210	0.00316	0.00343
% CV	7.2	6.9	8.4	7.2	8.8	7.6

Notes: The projected rates in this table are for purposes of illustration only. Tjaden and Thoennes (1997) obtained a rate of .003 for the prevalence of rape for females age 18 and over, and this value was used to forecast the results from the telephone sample for the general population. The age distribution for rape and attempted rape from the NCVS was used to extrapolate a rate of .003 for females age 18 and over to a rate, .0450, for females 18-49 for the in-person sample. Rates for other sexual assault were set at 7 times the rate for rape, roughly following ratios that have been observed in the British Crime Survey. Rates for the high risk sample were set at twice the rate for the in-person sample.

## B.2e Evaluation of In-Person and Telephone Designs

A primary goal of the analysis will be to evaluate how the approaches implemented on the NSHS improve measures of rape and sexual assault. The analysis will be guided by two basic questions:

1. What are the advantages and disadvantages of a two-stage screening approach with behavior-specific questions?
2. What are the advantages and disadvantages of an in-person ACASI collection when compared to an RDD CATI interview?

The analysis will examine the NSHS assault rates in several different ways. It is expected that comparisons to the NCVS will result in large differences with the NSHS estimates. We also will compare the estimates of the ACASI and telephone interviews. The direction of the difference in these estimates by mode may not be linked specifically to quality. If the rate for one mode is significantly higher than the other, it will not be clear which one is better. We will rely on a number of other quality measures, such as the extent to which the detailed incident form (DIF) improves classification of reports from the screener, the reliability of estimates as measured by the re-interview, the extent of coverage and non-response bias associated with the different modes, and the extent respondents are defining sexual assault differently. In the remainder of this response we review selected analyses to illustrate the approach and statistical power of key analyses.

### *Comparison of Assault Rates*

There will be two sets of comparisons of assault rates. One will be with the equivalent NCVS estimates and the second will be between the two different survey modes.

### Comparisons to the NCVS

A basic question is whether the estimates from NSHS, either the in-person or CATI approaches, differ from the current NCVS. The NCVS results are estimates of *incident rates*, that is, the estimated number of incidents divided by the estimated population at risk. However, the sources used in our design assumptions, such as the British Crime Survey, have emphasized lifetime or one-year *prevalence rates*, that is, the estimated number of persons victimized divided by the estimated population at risk. In general, prevalence rates for a given time interval cannot be larger than the corresponding incidence rates. For the NSHS, we expect to estimate both one-year prevalence and one-year incidence rates for rape and sexual assault. Our design assumptions for NSHS focus on one-year prevalence rates and we anticipate that the NSHS estimates for prevalence will be considerably larger than the incident rates estimated by the NCVS. This assumption is based on prior surveys using behavior-specific questions that

have observed rates that differ from the NCVS by factors of between 3 and 10, depending on the survey and the counting rules associated with series crimes on the NCVS (Rand and Rennison, 2005; Black et al., 2011). The use of ACASI and increased controls over privacy for both surveys has also been associated with increasing the reporting of these crimes (e.g., Mirrlees-Black, 1999). Thus, the NSHS design assumes that both the in-person and telephone approaches will yield estimates considerably larger than implied by the NCVS.

To assess statistical power, we derived an annual prevalence rate implied by the NCVS by computing the 6-month rate for women 18-49 years old and doubled it to approximate the 12-month rate. This clearly overestimates the actual rate because it excludes the possibility that some respondents could be victimized in both periods. Table 10 provides these approximate prevalence rates from 2005 to 2011. A multi-year average is used to remove fluctuations due to sampling error related to the small number of incidents reported on the survey for a particular year.

**Table 10. Approximate annual prevalence rates and standard errors for the NCVS for females 18-49 during 2005-2011.**

Year	Rape, attempted rape, and sexual assault		Rape and attempted rape only	
	NCVS estimate	Standard error	NCVS estimate	Standard error
2005	0.0014	0.0003	0.0011	0.0002
2006	0.0022	0.0004	0.0014	0.0003
2007	0.0025	0.0003	0.0015	0.0003
2008	0.0017	0.0003	0.0011	0.0003
2009	0.0012	0.0003	0.0009	0.0003
2010	0.0020	0.0004	0.0012	0.0003
2011	0.0022	0.0004	0.0016	0.0003
Pooled, 2005-2011	0.0020	0.0001	0.0013	0.0001
Pooled 2005 - 2011 for 5 metro areas	.0020	.0004	.0013	.0003

Notes: National rates and standard errors derived from the NCVS public use files, downloaded from ICPSR. Pooled rate for 5 metro areas are based on assumption that

the rate is the same as the national and the standard error is approximately 3 times as large as the national standard error.

During 2005-2011, the NCVS reflected an annual prevalence rate of approximately 0.0013 for rape and 0.0020 for rape and sexual assault combined (Table 10). Considering both the NCVS sample sizes in the five metropolitan areas and the effect of the reweighting to reflect the NSHS sample design, the standard errors for the NCVS for the five metropolitan areas combined are likely to be about 3 times as large as the corresponding national standard errors. Thus, the standard error for estimates from the NCVS for rape and attempted rape is likely to be about 0.0003 on the estimate of 0.0013, and the standard error for rape, attempted rape, and sexual assault about 0.0004 on the estimate of 0.0020 (Table 10).

Our design is based on the assumption that the NSHS prevalence rate for rape will be roughly 3 times the NCVS rate and a prevalence rate for rape and sexual assault roughly 15 times the NCVS rate based on the above studies. For design purposes, this translates to assumed prevalence rates for the NSHS of 0.0045 for rape (and attempted rape) and 0.0360 for rape and sexual assault. Table 11 provides the power of comparisons assuming that these differences occur.

**Table 11. Expected power for comparisons between expected NSHS Sexual Assault rates with the NCVS**

Survey	Rape, Attempted Rape and Sexual Assault			Rape and Attempted Rape		
	Estimate	Standard Error	Power+	Estimate	Standard Error	Power+
NCVS	.0013	.0003		.0020	.0009	
NSHS In-Person	.0045	.0009	90%	.036	.0026	>90%
NSHS: Telephone	.0045	.0011	80%	.036	.0032	>90%

+ Power when compared to the NCVS estimate

For the combined category of rape, attempted rape, and sexual assault, both the in-person and telephone samples easily will yield statistically significant findings when compared to the much lower NCVS results. Estimates of 0.0360 with standard errors of 0.0026 and 0.0032 for in-person and telephone, respectively, are certain to yield significant results when compared to 0.0020 from the NCVS. For the less frequent category of rape and attempted rape, for which an estimate of about 0.0045 is expected, the situation requires a closer check—the standard errors of 0.0009 and 0.0011 for in-person and telephone, respectively, would have associated power of about 90% in the first case and about 80% in the second.

Comparisons between the In-Person and CATI approaches

Table 12 provides the estimates of power for comparing the overall rates of rape and sexual assault between the two modes. To detect a significant difference in estimates of rape with 80% power, the estimates would have to differ by a factor of 2 (.004 vs. .008). While this large difference is not unusual for many of the comparisons discussed above, it is large when comparing two methodologies that are similar, at least with respect to the questionnaire. Estimates of other sexual assault will be able to detect differences of about 33% of the low estimate (e.g., .03 vs. .04). This drops to around 25% of the estimate when combined with the high risk sample (Data not shown). So there should be reasonable power for this aggregated analysis.



**Table 12. Size of the Actual Difference in Sexual Assault Rates Between Modes to Achieve 80% Power**

Type of Assault	Low Estimate in the Comparison	Standard Error of Difference between Modes	Size of Difference to have 80% Power
Rape	.0045	.00146	.044
Other Sexual Assault	.031	.0031	.0095

### *Use of a Detailed Incident Form*

As noted above, the comparison of the rates, while interesting, is not a direct measure of data quality. One of the primary goals of the NSHS is to develop and evaluate a detailed incident form (DIF) to classify and describe events (e.g., see research goal, question #1 above). Prior studies using behavior-specific questions have depended on the victimization screening items to classify an incident into a specific type of event. This methodology relies on the respondent's initial interpretation of the questions to do this classification. For the NCVS, this can be problematic because the screening section does not document the essential elements that define an event as a crime. For example, on the NCVS, a significant percentage of incidents that are reported on the screener do not get classified as a crime because they lack critical elements (e.g., threat for robbery; forced entry for burglary). We are not aware of the rate of 'unfounding' the Census Bureau finds from this process, but in our own experience with administering the NCVS procedures, 30% of the incidents with a DIF are not classified as a crime using NCVS definitions. Furthermore, the screener items may not be definitive of the type of event that occurred. For example, on the NCVS a significant number of events classified as robberies come from the initial questions that ask about property stolen, rather than those that ask about being threatened or attacked (Peytchev, et al., 2013).

The increased specificity of behaviorally-worded screening questions may reduce this misclassification. However, even in this case respondents may erroneously report events at a particular screening item because they believe it is relevant to the goal of the survey, but it may not fit the particular conditions of the questions. Fisher (2004) tested a detailed incident form with behavior-specific questions and found that the detailed incident questions resulted in a significant shift between the screener and the final classification. Similarly, our cognitive interviews found that some respondents were not sure how to respond when they experienced some type of sexual violation, but were unsure if it when asked a specific question. It may have been an alcohol-related or intimate partner-related event, which the respondent thought was relevant, but did not exactly fit when asked about 'physical force' (the first screener item). Some answered 'yes' to the

physical force question, not knowing there were subsequent questions targeted to their situation.

An important analysis for this study will be to assess the utility of a DIF when counting and classifying different types of events involving unwanted sexual activity. This will be done by examining how reports to the screener compare with their final classification once a DIF is completed. Initially, we will combine both in-person and CATI modes of interviewing for this analysis. This will address the question of whether a DIF, and its added burden, is important for estimating rape and other sexual assault. We will then test whether there are differences between the two different modes of interviewing. We will analyze the proportion of incidents identified in the screener as rape and sexual assault that the DIF reclassifies as not a crime, in other words, unbounding them.

**Table 13. Standard errors for the estimated unbounding rates, for true unbounding rates of .10 and .30.**

	In-Person		Telephone 18-49		Combined	
	General Populat'n	General Populat'n + high risk	General Populat'n	General Populat'n + high risk	General Populat'n	General Populat'n + high risk
Standard errors of estimated unbounding rate, for unbounding rate=0.10						
Rape	0.06197	0.05277	0.07682	0.06281	0.04823	0.04026
Other Sexual Assault	0.02342	0.01995	0.02904	0.02374	0.01823	0.01522
Total	0.02191	0.01866	0.02716	0.02221	0.01705	0.01423
Standard errors of estimated unbounding rate, for unbounding rate=0.30						
Rape	0.09466	0.08061	0.11735	0.09594	0.07368	0.06150
Other Sexual Assault	0.03578	0.03047	0.04435	0.03626	0.02785	0.02325
Total	0.03347	0.02850	0.04149	0.03392	0.02605	0.02174

Notes: The rates used in the calculations are taken from Table 9 of the October, 2013 submission to OMB. The rates are for purposes of illustration only. The standard errors shown are the standard errors of the estimated unbounding rate given the crime rate.

The results for the illustrative unfounding rates of 0.10 and 0.30 are shown in Table 13. These show that the sample sizes will allow for assessing the overall value of the DIF. For all sexual assaults the confidence intervals for a rate of 10% will be  $\pm 0.03$  or less. For example, if the estimate of unfounding is 10%, the study would estimate that the use of the DIF would reduce the rates implied by the screener between 7% and 13%. If the unfounding rate is as high as 30%, then the use of a DIF will reduce the rate from the screener by 25% to 35%. This should provide the needed perspective on the relative merits of the DIF for classifying events as crimes.

Comparisons of the two modes with respect to relatively low unfounding rates will be able to detect differences of about 10 percentage points with 80% power for all sexual assaults. For example, there is 80% power if one mode has an unfounding rate of 5% and the other a rate of 15%. If the unfounding rates are higher, true differences of around 16 percentage points is required (e.g., 25% vs. 41%) for 80% power.

The confidence intervals for the unfounding estimate for rape will be much broader. If the rate is around 10%, the confidence intervals will be as large as the estimate. Higher rates (e.g. around 30%), which are not likely given the relatively small number reported, will be  $\pm 15\%$ . For example, if the estimate is 30%, the confidence interval will be between 15% and 45%.

### *Re-interview and Reliability*

A second measure of quality will be estimates of reliability from the 1,000 re-interviews that will be conducted with those who report some type of unwanted sexual activity on the screener. The goal of the re-interview will be to assess simple response variance (e.g., consistency or reliability of the response data; e.g., Biemer et al, 1991: 16-21). An important element of quality is whether respondents interpret the questions the same way once they are exposed to the entire interview. On the NCVS, for example, there are significant effects associated with repeated interviewing (e.g., Biderman and Cantor, 1984). The extent that respondents change answers is indicative of problems with the instrument. For example, an important question will be whether respondents either do not report an incident at the

re-interview or if the incident changes classification. Similarly, change in the dating of the event or which screener items elicited the event provides indications of the stability of estimates.

The followup interview will use the same mode and instrument as the initial interview. The analysis will rely on several difference measures of consistency, including gross and net difference rates and Cohen's Kappa. To illustrate the power, using Kappa, Table 14 provides estimates of standard errors for a Kappa statistics when there are 10% and 20% inconsistencies for each of the sample groups. For the general population sample, comparing Kappa's with these levels of inconsistencies will have 50% power to detect an actual difference of approximately 10 percentage points. There would be 80% power to detect a difference of 15 percentage points. The differences would have to be slightly larger for the high risk respondents. It will also be possible to combine the high risk and general populations into a single estimate. This will reduce the standard errors to some degree, although will also limit the generalizability of the estimate.

There will be higher power for simpler and more descriptive comparisons that do not involve computing Kappa statistics. Some of these will lead to higher power. For example, one analysis will compare the simple percentage of those changing classifications. The power of this analysis for those reporting victimizations during the previous year would be almost 75% when comparing a 10% to a 20% reclassification rate.

Table 14. Estimated Values and Precision of Kappa Statistics by sampled group

	<b>Kappa 1 (10%)</b>		<b>Kappa 2 (20%)</b>	
	<b>Value</b>	<b>S.E.</b>	<b>Value</b>	<b>S.E.</b>
ACASI, general	0.881	0.039	0.763	0.051
ACASI, high risk	0.863	0.046	0.723	0.063
CATI, general	0.881	0.033	0.763	0.043
CATI, high risk	0.863	0.046	0.723	0.063

### **Coverage and Non-Response of Important Subgroups**

An important consideration when choosing between an in-person ACASI and an RDD telephone interview is the extent to which there are relative differences in coverage and non-response at the unit level. Differences in non-response are especially important given the continued drop in response rates for RDD surveys. If the differentials are for groups that are thought to be high risk, such as college students or low income groups, there will be a difference in how well the two modes are representing the target population.

The NSHS design will not allow clean separation of the effects of coverage from non-response. However, by comparing the distributions of the final set of respondents, it will be possible to assess their combined effects. If there are differences in the distributions, data from the ACS can be used to assess which survey was closer to the truth.

Table 15 provides the size of the difference needed to achieve 80% power. With 7500 and 4880<sup>3</sup> interviews in the ACASI and telephone samples respectively, the power for these comparisons will be quite high. For example, estimates as low as 1% in the population will have a standard error for a difference of around .22%. Comparisons of around this percentage would have 80% power when the actual difference was around .7%. As can be seen, proportionately smaller differences would be needed to have 80% power to detect a statistically significant difference.

---

<sup>3</sup> There will be 4880 respondents age 18-49 for the RDD CATI survey.

**Table 15. Actual Difference for 80% Power when comparing estimates of Respondent Characteristics**

% in Sample	Standard Error of Difference	Actual Difference required for 80% power
1%	0.22%	0.7%
5%	0.48%	1.4%
10%	0.66%	2.0%
20%	0.88%	2.6%
30%	1.01%	3.0%
40%	1.08%	3.2%
50%	1.10%	3.3%

+Assumes a design effect of 1.4 for both the ACASI and Telephone surveys.

### *Respondent Interpretation of in-scope incidents*

The two designs will be compared on how they differ with respect to respondent comprehension of key concepts in the questionnaire. This comparison will be indicative of whether the two designs differ in how respondents may be interpreting the rape and sexual assault questions. At the end of the survey, each respondent will be asked how she would answer the selected screener questions for different scenarios. The scenarios are structured to vary by key dimensions, such as the level of coercion, the type of (non)consent provided and whether alcohol was involved in the situation.

Observed differences between modes could be due to a number of factors, including differential non-response or mode effects (among others).

Regardless of the reasons for the differences, this will provide perspective on whether the two designs are leading to noticeably different interpretations of key concepts.

For example, in one scenario, respondents will get one of four different situations related to consent:

1. Tom asked if she wanted to have sex. Sue said yes and they proceeded to have sexual intercourse.
2. Tom kissed Sue and they proceeded to have sexual intercourse. Sue did not say anything at the time, but she did not want to have sex.
3. Tom kissed Sue. She tried to push Tom away, but did not actually say no. They proceeded to have sexual intercourse.

4. Tom kissed Sue. Sue said she did not want to have sex, but Tom ignored her and they proceeded to have sexual intercourse. Sue did not resist again because she was afraid Tom would hurt her.

Respondents will be randomly assigned to one of these four conditions. After reading the randomized vignette, respondents are asked a few questions, four after the first vignette and two after the second. Responses to each of the questions can be analyzed separately. The simple analysis will examine the extent the ACASI and telephone responses differ. For example, it will compare the percent of ACASI respondents who report the scenario as a rape when condition 3 is used to those who report it when condition 4 is used.

For purposes of comparing modes, we will restrict the analysis to the ACASI sample and telephone respondents age 18-49, with target sizes of 7,500 and an expected 4,880, respectively. The two modes can also be combined into a sample of over 12,000 for detailed analysis. Because of the randomization, respondents will only receive one from the set of situations, where the number of situations varies from two (involving work) to five (involving drinking behavior and involving relationship). The total sample size is divided among the situations along each dimension. For example, to use 12,000 as the approximate the overall sample size, 3,000 respondents will have answered each of the situations above.

With these sample sizes, the power to detect relatively small differences between modes will be quite high. For example, the third item above, there would be approximately 1,875 ACASI and 1,220 telephone respondents. For example, if the question received a 20% "yes" response, the standard errors of the estimates would be about .009 for ACASI and .011 for telephone. The power to detect differences of 4.4% between the modes would be about 80%. We will also conduct more detailed multivariate analyses that examine differences between modes once accounting for other experimental conditions, such as the respondent's prior relationship. This should increase the overall precision of the estimates.



## **B.3 Methods to Maximize Response**

### **B.3a Address Based Sample**

Westat will recruit a team of experienced field interviewers and field supervisors sufficient in size to work all cases thoroughly. These field staff will be strategically located within or in close proximity to PSUs, which will expedite visits to the sample dwelling units and will also ensure that they are familiar with the communities within which the cases are located. Field interviewers will also be thoroughly trained in gaining respondent cooperation through refusal aversion and conversion. Field management will ensure that data collection efforts are thoroughly planned down to the field interviewer level; for example, production goals will be developed that will set a pace for individual field interviewers, field supervisor teams, and across all cities in this pilot study.

Several tools and approaches to address nonresponse and maximize response rates will be used. First, the interviews will be conducted in Spanish in addition to English and bilingual field staff will be available to administer them. Second, respondent materials will be developed to encourage participation, also translated into Spanish. These materials will include an advance letter to inform selected households of the study prior to in-person contact. The advance letter contains assurances of privacy and describes the voluntary nature of the survey and principal purposes and uses of the survey data. A toll-free call line to answer respondents' questions and to reassure them of the credibility of the study will be established. Tailored letters will be developed for use with reluctant respondents/sample persons and with selected units located in limited-access situations (doorperson buildings, gated communities, etc.), which may be sent via FedEx or priority mail to reinforce the perceived importance of participation. (Attachment R)

A web-based Supervisor Management System (SMS) will allow field supervisors to closely monitor each field interviewer's work, which facilitates the development of strategies to address nonresponse. These strategies will include reassigning difficult or reluctant cases among local field interviewers

and the use of specially trained, traveling field interviewers who are highly skilled in refusal conversion.

In addition to these measures, respondents will be offered an incentive to participate. The rationale for this is provided in section A.9.

### **B.3b Telephone Procedures to Maximize Response**

To increase response to the telephone survey (feasibility and pilot), interviewers with experience conducting RDD surveys will be used to the extent possible, and multiple callbacks will be made using an automated call scheduling system. Interviewer training will focus on gaining cooperation in the first minute or so of the initial contact with a potential respondent. To maximize the contact rate, we will use a calling algorithm that handles all dimensions of call scheduling, including time zone (respondent and interviewer); skill level of interviewer, special needs of the case (e.g., non-English language); call history; and priority of case handling.

The data collection design includes up to 14 call attempts to determine whether the telephone number reaches a household; if there are 14 non-contacts (across a variety of times and days) then the sample record will be closed as final nonresponse. Those telephone numbers determined to belong to a household will be called until we obtain a completed interview or until the household refuses to participate. Those respondents deemed to be hostile to the survey request will be coded out as a final refusal. For the non-hostile respondents who refuse to participate, a specially-trained interviewer will recontact and attempt to convert the refusal to an interview. Those respondents who refuse twice will be coded as final nonresponse. Refusal conversions will be attempted at both screening and main interview stages. Interviews will be conducted in English and Spanish languages.

For the landline sample, advance letters will be mailed for sampled numbers matched to addresses, to announce the study and share information about the purpose and sponsor of the study. This type of advance mailing is presently only possible for the landline sample, as commercial databases of

addresses for cell phone numbers are not readily available or reliable at this time. In addition, refusal conversion letters will be mailed to landline households with addresses prior to attempting conversion.

Review sessions will be held with the interviewing staff roughly two weeks after the start of data collection. During these sessions, common obstacles to gaining cooperation for this study will be reviewed and techniques or tools the data collectors are employing to surmount these obstacles will be shared.

### **B.3c Calculation of Response Rates**

**Address sample.** Because the design is relatively straightforward, weighted and unweighted response rates for the address sample can be determined in the standard manner. Base weights will be computed for all of the units in the designated sample.

Addresses will be classified as occupied, vacant, or not currently eligible to be a dwelling or group quarters unit. A failure-to-locate rate will represent situations where interviewers are unable to establish the location of ABS sample addresses. Because the design includes a coverage improvement operation, occupied ABS addresses that cannot be located should in theory be represented by coverage additions from the missing address procedure, so a small number of units that cannot be located would not significantly impact the overall quality of the study.

Another contributor to nonresponse will be units where the interviewer is unable to determine the occupancy status of a sampled unit. For example, lack of access to some gated communities may prevent completing the classification of the entire sample. A separate rate will be computed to represent this component of nonresponse.

The response rate to the screener will be defined as the number of households completing the screening interview as a percent of sampled occupied units. If the occupancy status of some addresses cannot be

determined, then the denominator should be the estimated number of occupied units, where the occupancy status for units that cannot be classified is estimated based on the rest of the sample.

The screening interview will be counted as complete if the interviewer obtains a complete roster, enabling subsampling of an eligible female, if any are present. The denominator of this rate will be the number of occupied units.

Similarly, a non-response rate can be calculated for the ACASI interview. The unweighted and weighted sum of completed screening cases will be compared to the unweighted and weighted sum of households completing the screening interview. These rates will be reported for each CBSA separately.

**Telephone sample.** For each of the telephone samples both unweighted and weighted screener response rates will be calculated separately for each CBSA. To calculate the weighted screener response rate, the “base” weight reflecting the probability of selecting the telephone household will be used. For example, the weighted screener response rates will be calculated by taking the sum of weights for the set of households in the CBSA that responded to the screening interview divided by the sum of weights for the set of households in the stratum that are estimated to be residential. That is, the weighted screener response rate, SRR will be computed from the general formula:

$$SRR = R / (R + N + e(U)) \quad (1)$$

where R = the weighted number of screener respondents, N = the weighted number of screener nonrespondents known to be residential, and e(U) is an estimate of the residential cases among the unknown-eligibility cases.

The next set of response rates will be at the extended interview level. The extended interview response rate for a CBSA is the weighted percentage of the women sampled in the screener interview who completed an extended

interview. The weight used in this calculation is the inverse of the probability of selecting the women in the household. The extended interview response rate, IRR, is computed as the weighted number of responding women divided by all eligible women. The overall response rate, ORR, for each CBSA within a sample will be the product of the screener and extended interview response rates; i.e.,  $ORR = SRR * IRR$ .

We can also compute the overall response rate for the combined landline RDD and cell phone samples. The adult response rate is the weighted average of the overall response rates of women in the landline and cell only. These overall rates are weighted by the female population by type of telephone service (i.e., cell-only, both, landline-only telephone services). The overall combined response rate for the landline and cell phone sample,  $ORR_{comb}$ , is computed as

$$ORR_{comb} = p_{cell\_only} \cdot ORR_{cell\_only} + p_{both} \cdot ORR_{both} + p_{landline\_only} \cdot ORR_{landline\_only} \quad (2)$$

where  $ORR_{cell\_only}$  is the overall response rate of cell-only women sampled in the cell phone sample;  $ORR_{both}$  is the overall response rate of women who have a cell phone and a landline from either cell phone and landline samples; and  $ORR_{landline\_only}$  is the overall response rate of landline-only women from the landline sample. The proportions  $p_{cell\_only}$ ,  $p_{both}$ , and  $p_{landline\_only}$ , are the proportions by type of telephone service that are shown in Table 2.

The overall cell-only adult response rates  $ORR_{cell\_only}$  and  $ORR_{landline\_only}$  are computed as the product of the corresponding screener response rate (i.e., cell or landline screener response rates) and the corresponding adult extended response rate (i.e., cell-only adult extended interview rates or landline-only adult extended interview response rates). In contrast, since the overall adult response rate with both telephone services,  $ORR_{both}$ , combines samples from the cell phone and landline samples, this response rate is computed as the weighted average of the overall response rates of adults with both telephone services from the two samples as

$$ORR_{both} = \lambda \cdot ORR_{both}^{landline\_smp} + (1 - \lambda) \cdot ORR_{both}^{cell\_smp} \quad (3)$$

where  $ORR_{both}^{landline\_smp}$  and  $ORR_{both}^{cell\_smp}$  are the overall response rates for adults with both types of telephone in the landline and cell phone samples respectively and  $\lambda$  is the composite factor (see section B.2c )used to combine these rates form the two samples.

## **B4. Test of Procedures or Methods**

### **B4a. Cognitive Testing**

In February/March 2013 and in May 2013, two rounds of cognitive interviews were conducted to test and refine the ACASI and CATI instruments that will eventually be administered in the feasibility study and pilot test. Round 1 consisted of 23 women recruited as high risk sample from Craig’s List and local colleges and universities in the Washington, DC area. Westat recruiters screened the women to identify whether they had ever had any unwanted sexual experiences; 18 of the 23 women who ultimately participated in the interviews had experienced some unwanted sexual act in their lifetime, including five in the past 12 months. All respondents were asked to be re-interviewed two weeks later; 22 out of 23 women completed the re-interview. Additionally, 9 women were recruited from the DC Rape Crisis Center to participate in a cognitive interview as Service Provider sample. Key findings from the first round of interviews showed that the detailed incident form required greater refinements in how it asks about the types of assaults that were threatened, attempted, or completed; that greater detail was needed on the role of alcohol and drugs in the incident; and that de-duplication efforts in the screening questions were often confused by asking about lifetime estimates interleaved with past 12 month estimates.

Round 2, conducted in May 2013, consisted of 20 women who had had unwanted sexual experiences in their lifetime, with an emphasis on those having an experience in the past 12 months. Women were recruited from Craig’s List and local colleges and universities in the Washington, DC area.

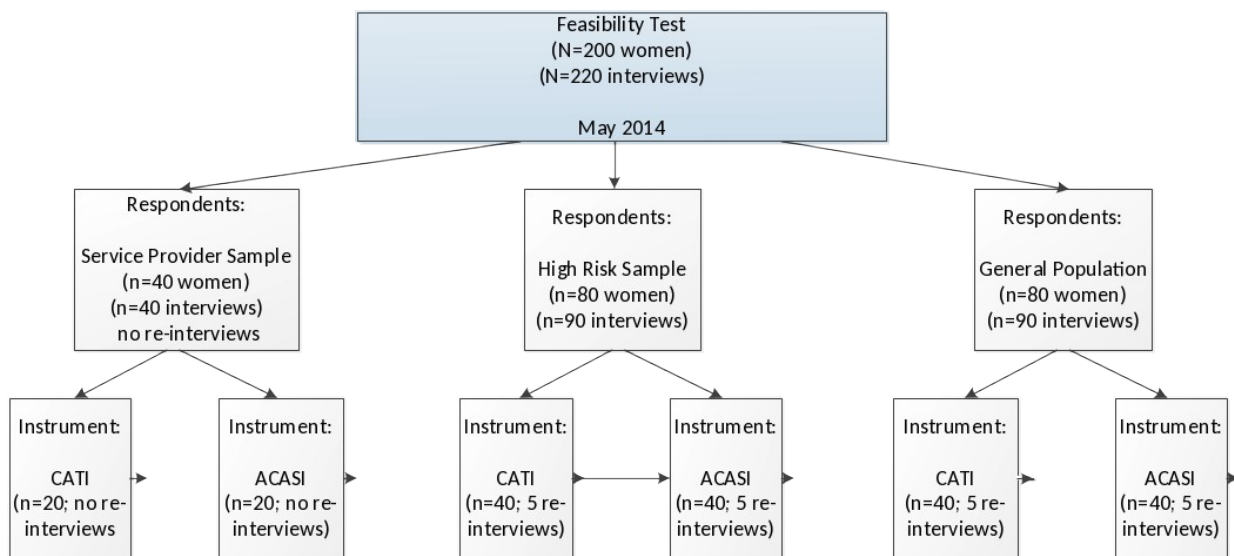
Results from these interviews will aid in the final refinement of the instrument.

Spanish cognitive interviewing were conducted in October, 2013 after translation of the final instruments, allowing for a final set of adjustments to ensure the Spanish-speaking population understands the questions as intended.

**B4b. Feasibility study**

Over a 4-week period, Westat will recruit a total of 200 respondents across the same core based statistical areas, or CBSAs (Los Angeles, Phoenix, Dallas, Miami, and New York City) for a small feasibility study to test field operations and identify any outstanding issues prior to the pilot study. Respondents will be recruited from the same types of samples as for the pilot test: general population, respondents at a high risk for sexual assault, and a service provider sample recruited from local rape crisis centers. Both ACASI and computer assisted telephone interview (CATI) administration modes will be used for each sample.

Figure 2. Project Number of Completed Interviews by Sample Components for NSHS Feasibility Survey



Based on analysis of the results of the feasibility test, last minute adjustments will be proposed to BJS in June and will be submitted to OMB for consideration and approval 6 weeks after the end of the field period.

**B5.       Contacts for Statistical Aspects and Data Collection**

BJS contacts are:

Allen J. Beck, Ph.D.  
Senior Statistical Advisor  
Bureau of Justice Statistics  
810 Seventh St., N.W.  
Washington, DC 20531  
(202) 616-3277



Shannan Catalano, Ph.D.  
Statistician  
Bureau of Justice Statistics  
810 Seventh St., N.W.  
Washington, DC 20531  
(202) 616-3502

Westat contacts are:

David Cantor, Ph.D.  
Principal Investigator  
1600 Research Blvd.  
Rockville, MD 20850  
(301) 294-2080

Robert Fay, Ph.D.  
Senior Statistician  
1600 Research Blvd  
Rockville, MD 20850  
(240) 314-2318

## References:

- Bialik, C (2013) "Statistics Shed Little Light on Rape Rate." *Wall Street Journal*, August 25, 2013.
- Biderman, A.D., and Cantor, D. (1984). A longitudinal analysis of bounding, respondent conditioning and mobility as sources of panel bias in the National Crime Survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 708-713.
- Biderman, A.D., Cantor, D., Lynch, J.P., and Martin, E. (1985). *Final report of the National Crime Survey redesign*. Washington, DC: Bureau of Social Science Research.
- Biemer, P.B., Groves, R.M., Lyberg, L., E., Mathiowetz, N.A., and Sudman, S. (1991). *Measurement Error in Surveys*; Wiley Series in Probability and Mathematical Statistics.
- Black, M.C., Basile, K.C., Breiding, M.J., Smith, S.G., Walters, M.L., Merrick, M.T., Chen, J., & Stevens, M.R. (2011). The National Intimate Partner and Sexual Violence Survey (NISVS): 2010 Summary Report. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.
- Black, M.C., Kresnow, M., Simon, T., Arias, I., & Shelley, G. (2006). Telephone survey respondents' reactions to questions regarding interpersonal violence. *Victims and Violence*, 21(4), 445-459.
- Brackstone, G.J., and Rao, J.N.K. (1979). An investigation of raking ratio estimation. *Sankhya*, 41, 97-114.
- Bradburn, N. M. (1978). Respondent Burden. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 35-40.
- Brick, J. M. and D. Williams (2013) "Reasons for increasing nonresponse in U.S. household surveys." *Annals of the American Academy of Political and Social Science*. 645: 36-59.
- Brick, J. M., Flores Cervantes, I., Lee, I., and Norman, G (2011). Nonsampling errors in dual frame telephone surveys. *Survey Methodology*, June 2011, Vol. 37, No. 1, pp. 1-12, Statistics Canada.
- Brick, J. M., Montaquila, J., Hagedorn, M.C., Roth, S.B., and Chapman, C. (2005). Implications for RDD Design from an Incentive Experiment. *Journal of Official Statistics*. Vol. 21, No. 4, pp. 571-589.

- Brick, J.M., Waksberg, J., Kulp, D., and Starer, A. (1995). Bias in list assisted telephone surveys. *Public Opinion Quarterly*, 59(2), 218-235.
- Campbell, R., & Adams, A.E. (2009). Why do rape survivors volunteer for face-to-face interviews? A meta-study of victims' reasons for and concerns about research participation. *Journal of Interpersonal Violence*, 24(3):395-405.
- Cantor, David, Holly Schiffrin, Inho Park and Brad Hesse. (2006) "An Experiment Testing a Promised Incentive for a Random Digit Dial Survey." Presented at the *Annual Meeting of the American Association for Public Opinion Research*, May 18 - 21. Montreal, CA.
- Cantor., D., O'Hare, B. and O'Connor, K. (2008) "The Use of Monetary Incentives to Reduce Non-Response in Random Digit Dial Telephone Surveys" pp. 471-498 in J. M. Lepkowski, C. Tucker, J. M. Brick De Leeuw, E., Japec, L., Lavrakas, P. J., Link, M. W., & Sangster, R. L. (Eds.), *Advances In Telephone Survey Methodology*, New York: J.W. Wiley and Sons, Inc.
- Cantor, D. and D. Williams (2013) Assessing Interactive Voice Response for the National Crime Victimization Survey. Final Report prepared for the Bureau of Justice Statistics under contract 2008-BJ-CX-K066
- Church, Allan H. 1993. "Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis." *Public Opinion Quarterly* 57:62-79.
- Curtin, Richard, Stanley Presser, and Eleanor Singer. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly* 69 (1): 87-98.
- Edwards, P., Cooper, R., Roberts, I. and C. Frost (2005) "Meta-analysis of randomized trials of monetary incentives and response to mailed questionnaires." *Journal of Epidemiology and Community Health* 59: 987-999.
- Fisher, B. (2004). *Measuring rape against women: The significance of survey questions*. NCJ 199705, US Department of Justice, Washington DC.
- Fisher, B.S., & Cullen, F.T. (2000). Measuring the sexual victimization of women: Evolution, current controversies, and future research. *Criminal Justice*, 4, 317-390.
- Gilbert, N. (1997) "Advocacy research and social policy." In Tonry, M. (ed) *Crime and justice: A review of research*. Volume 22. Chicago: University of Chicago Press

- Hall, P. and K. Smith (eds) (2011) "Analysis of the 2010/11 British Crime Survey intimate personal violence split sample experiment." Crime Statistics Programme, Home Office Statistics Unit, Home Office Science Group.  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/116718/analysis-bcs-ipv-2011.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/116718/analysis-bcs-ipv-2011.pdf)
- Hartley H. O. (1962). Multiple-frame surveys. Proceedings of the Social Statistics Section. American Statistical Association, 203-206.
- Holt, D., and Smith, T.M.F. (1979). Post stratification. *Journal of the Royal Statistical Society. Series A (General)*, 142(1), 33-46.
- Kilpatrick, D., Resnick, H., Rugiero, K., Conoscenti, L., and McCauley, J. (2007) *Drug-Facilitated, Incapacitated, and Forcible Rape: A National Study*. Charleston, SC: Medical University of South Carolina and National Crime Victims Research and Treatment Center.
- Koss, M.P., and Gidycz, C. (1985). Sexual experiences survey: Reliability and validity. *Journal of Consulting and Clinical Psychology*, 53(3), 422-423.
- Kreuter, F., Presser, S. and R. Tourangeau (2008) "Social Desirability Bias in CATI, IVR and Web Surveys. The Effects of Mode and Question Sensitivity." *Public Opinion Quarterly* 60: 410-430.
- Labott, S.M., Johnson, T., Fendrich, M. and N.C. Feeny (2013) Emotional Risks to Respondents in Survey Research: Some Empirical Evidence. *Journal of Empirical Research on Human Research Ethics: An International Journal*, Vol. 8, No. 4: 53-66.
- Lynch, J.P. (1996). Clarifying Divergent Estimates of Rape from Two National Surveys." *Public Opinion Quarterly* 60: 410-430.
- Mirrlees-Black, C. (1999). Domestic Violence: Findings from a new British Crime Survey self-completion questionnaire. Home Office Research Study 191, London.
- Peytchev, A., Caspar, R., Neely, B. and A. Moore (2013) *NCVS Screening Questions Evaluation: Final Report*. Research Triangle Institute, Research Triangle Park, NC.
- Rand, M.R, and Rennison, C.M. (2005). Bigger is not necessarily better: An analysis of violence against women estimates from the National Crime Victimization Survey and the National Violence Against Women Survey. *Journal of Quantitative Criminology*, 21, 267-291.
- Research Triangle Institute. (2002). 2001 National Household Survey on Drug Abuse: Incentive Experiment. Rockville, Maryland: Project 7190 - 1999-2003 NHSDA.

- Singer, E. (2002). The use of incentives to reduce nonresponse in household surveys. In R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little (Eds.), *Nonresponse in household interview surveys* (pp. 163-178). New York: John Wiley and Sons.
- Singer, Eleanor, John Van Hoewyk, Nancy Gebler, Trivellore Raghunathan, and Katherine McGonagle. (1999). "The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys". *Journal of Official Statistics* 15:217-230.
- Strouse, Richard C., and John W. Hall. (1997). "Incentives in Population Based Health Surveys". *Proceedings of the American Statistical Association, Survey Research Section*: 952-957.
- Tjaeden, P., and Thoennes, N. (1998). *Stalking in America: Findings from the National Violence Against Women Survey*. Research in Brief, NCJ 169592. Washington DC: U.S. Department of Justice.
- Tjaeden, P., and Thoennes, N. (2000). *Full report of the prevalence, incidence, and consequences of violence against women: Findings from the National Violence Against Women Survey*. Research Report NCJ 183781. U.S. Department of Justice, Office of Justice Programs.
- Tourangeau and Yan (2007) "Sensitive Questions in Surveys" *Psychological Bulletin* 133(5): 859-883.
- Truman, J. (2011) *Criminal Victimization, 2010*. September 2011. Bureau of Justice Statistics, US Department of Justice, NCJ 235508.
- Trussell, Norman, and P. Lavrakas. 2004. "The influence of incremental increases in token cash incentives on mail survey response: Is there an optimal amount?" *Public Opinion Quarterly* 68, no. 3: 349-367
- Tucker, C., Lepkowski, J., and Piekarski, L. (2002). The current efficiency of list assisted telephone sampling designs. *Public Opinion Quarterly*, 66, 321-338.
- Walker, E.A., Newman, E., Koss, M. and D. Bernstein (1997) "Does the study of victimization revictimize the victims?" *General Hospital Psychiatry* 19: 403-410.