Multiple Frame Approaches to Identify and Survey Victims of Rape and Sexual Assault[1]

James R. Chromy and David Wilson

RTI International

1. Introduction

Multiple frame methodology can be used to improve survey coverage and/or to reduce cost while maintaining broader coverage. This paper explores possible application to obtaining data about victims of rape and sexual assault. It first describes some current total population surveys which identify victims of rape and sexual assault while collecting data on a broader range of victimizations. These surveys have a low yield of victims of rape and sexual assault. The existence of some additional lists of known victims or rape and sexual assault is discussed and assumed for the purpose of applying multiple frame methodology.

2. Current Data Sources

There are two sources of national estimates of rape and sexual assault; the National Crime Victimization Survey (NCVS, 2010) and the National Intimate Partner and Sexual Violence Survey (NISVS, Black et al 2011). Rape and sexual assault are now measured in the NCVS administered by the U.S. Census Bureau on behalf o f the Bureau of Justice Statistics. The NCVS collects victimization data via a survey of a nationally representative sample of residential addresses and has been conducted since 1972. The NCVS collects victimization data for the six months prior to interview date.

The CDC's NISVS also measures rape and sexual assault. The NISVS collects data on past-year and lifetime experiences via a nationally representative telephone survey of the non-institutionalized English or Spanish-speaking adult population and was first conducted in 2010.

National Crime Victimization Survey

The NCVS covers all persons in the United States aged 12 or over. Using a rotating panel, households are randomly selected each month and all individuals 12 or older become part of the panel.

[1] "This paper was commissioned for the National Academies Workshop on Measuring Rape and Sexual Assault in BJS Surveys. The workshop was convened by the Committee on National Statistics on June 5-7, 2012 in Washington, DC with support from the Bureau of Justice Statistics.

Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted or endorsed or verified as accurate by the Committee on National Statistics or the National Academy of Sciences, including the National Academy of Engineering, Institute of Medicine, or National Research Council."

Respondents are interviewed every six months for a total of seven interviews over three years. Face to face interviews are conducted for the first and fifth interviews while the remaining interviews are conducted by telephone.

The households in the NCVS are selected using a stratified, multi-stage cluster design. The primary sampling units (PSUs) were geographic areas: counties, groups of counties, or large metropolitan areas. Large PSUs were selected with certainty while the remaining PSUs were assigned to sampling strata based on geographic and demographic characteristics calculated from the decennial Census. Additional stages of sampling are used, within PSU, to select housing units and other living quarters (e.g. group quarters).

In 2010, 106,071 households and approximately 167,000 persons were interviewed and the response rates were 92% and 88%, respectively.

National Intimate Partner and Sexual Violence Survey

The NISVS covers all adults in the United States and is conducted using a national random digit dial (RDD) telephone survey that utilizes a dual-frame approach that includes landline and cell phones. The 2010 survey was conducted in the 50 states and the District of Columbia and resulted in approximately 18,000 respondents (or partial respondents) and the weighted response rate varied from 28% to 34% (depending upon variation in estimates of unknown eligibility).

The NISVS sample was selected by stratifying by state in order to allow for state-level and national-level estimates. Eligible persons answering a cell phone were selected into the sample while an eligible person answering a landline phone was selected into the sample if the corresponding household was a one-adult household. If the household associated with a landline phone was a two-adult household, one adult was randomly selected into the sample. If the household associated with a landline phone consisted of more than two adults, the adult with the most recent birthday was selected into the sample.

In addition to the general NISVS sample, supplemental samples were drawn from two other populations: 1) persons of American Indian or Alaska Native ethnicity and 2) female active duty military and female spouses of active duty military2.

Survey and Sample Design Characteristics

While both the NCVS and NISVS may be used to generate national estimates of rape and sexual assault, there are differences between the two surveys; for example, their target populations, sampling frames, sampling design, and data collection modes are different. Various characteristics of the NCVS and NISVS surveys are listed and described in Table 1.

Aside from differences in target population, sampling frames, sampling design, and data collection mode; note that, everything else being constant, the number of interviews required to find one female victim of rape and sexual assault is, roughly, 1,674 for the NCVS and 91 for the NISVS. As will be illustrated, the degree to which supplemental sample frames are useful depends upon, among other things,

[2] Data not yet available for these two subpopulations.

the cost per completed interview of rape and sexual assault victims associated with the primary sample frame. Given the different costs per completed interview of victims, supplemental sample frames that would benefit the NCVS may be different from sample frames that benefit the NISVS or the benefit of existing supplemental frames may be different for the NCVS than the NISVS.

While both the NCVS and NISVS report estimates of rates of rape and sexual assault, the estimates can be quite different. Using the most recent data available for each survey as of June 2012, NISVS estimates that 1.3 million adult women were victims of rape in the year preceding the 2010 survey while NCVS estimates approximately 169,000 women aged 12 or older were the victims of rape or other sexual assault in 2010. Even though the two surveys use different age cutoffs in their population definitions and the associated retrospective time periods are not identical, the large difference in the estimates of the number of female victims is notable.

At the time of the writing of this paper, standard errors of these two estimates were not available but some rough estimates of the standard errors were made in order to assess the difference in totals in terms of precision. For the NCVS, the standard error of the total number of female and male victims of rape and sexual assault in 2010 is 29,399 (Truman 2010) and since the majority of reported victims are female, we can use 30,000 as an approximate standard error of the total number of female victims. Under the assumption of simple random sampling, the standard error of the NISVS estimate of the number of female victims of rape is, approximately, 143,000. Given these two standard error approximations, adding two standard errors to the NCVS estimate and subtracting two standard errors from the NISVS estimate gives values of 229,000 and 1,014,000. While these calculations are only approximate, there is evidence to suggest that the NCVS and NISVS total estimates are statistically different.

The total estimates for the NCVS and the NISVS are both based on relatively few victims so both surveys may benefit from introduction of supplemental frames in order to increase the number of cases used to estimate national totals.

Table 1. Two Primary Sources of National Estimates on Rape and Sexual Assault

Characteristic\Survey	**NCVS**	**NISVS**
Population	All persons 12 or older in the U.S. living in households and group quarters	All persons 18 or older in the U.S.
Design	Household sampling combined with rotating panel	Random-digit dialing
Collection Mode	Telephone, computer-aided telephone, and face to face	Telephone
Anchor Point	Date of survey	Date of survey
Recall Period	6 months (from first day of month of survey date)	12 months
# 2010 Interviews	~160,000	~18,000
# Female Interviews	~87,000	~9,100
# Female Rape/Sexual Assault Victims	52	~100

Estimated # of Female Victims of Rape and Sexual Assault[3]	~169,000	~1,300,000

Other Sources of Rape and Sexual Assault Victims

Both the NCVS and NISVS may benefit from using supplemental frames to improve national estimates of rape and sexual assault crimes. While the NCVS and NISVS provide the only, current, national estimates of rape and sexual assault, there are other sources of information about rape and sexual assault victims that might be able to be used to construct supplemental frames. Some of these other sources include emergency rooms, health clinics, insurance companies, and courts. While other sources exist, there may be barriers to access that include: privacy laws and other legal constraints, general privacy concerns, police enforcement, informed consent, and financial costs. For the remainder of this paper, we presuppose that supplemental frames could be constructed though we acknowledge that significant barriers may preclude constructing such frames in practice.

3. Population Definition

This paper is about multiple (sampling) frames, but it is important to recognize the conceptual population. The statistics produced from the survey or other data collection operation should reasonably apply to the conceptual population. The population needs to be defined geographically and temporally. Human populations may also be defined by demographics. The temporal definition is particularly important if events are to be measured.

An example of a population definition that could apply to victims of rape and sexual assault is:

- Female victims of rape and sexual assault
- 12 years old or older.
- Residing in the United States (50 states plus the District of Columbia) during any part of a designated year when the victimization occurred.

Note that this conceptual definition does limit the population to persons residing in households or to persons who are currently alive. Decisions about including (or excluding) persons currently residing in institutions or members of the military services need to be incorporated into the conceptual population definition. A distinction must also be made between the victimization events and the victims. The process of specifying the conceptual study population begins to suggest a variety of sources for

[3] The NISVS estimate is reported as an estimate of the number of female rape victims in the 12 months prior to the 2010 NISVS survey while the NCVS estimate is reported as an estimate of the number of female victims of rape and sexual assault in 2010.

identifying a sample of victims in some cases or a sample of victimizations in other cases. A different strategy may be required to estimate the number of victims vs. the number of victimizations.

4. Analysis Plans

One of the first steps in conducting a sample survey is the construction of an analysis plan. This plan defines the population(s) of interest and the research questions and analysis goals of the survey. Furthermore, this plan describes precision requirements needed to achieve the analytic goals of the survey. The analysis plan informs the choice, selection, and use of sampling frames.

For this paper, we focus on a simple analysis goal of estimating totals and note that survey data allow for unbiased estimates of totals where inference is to the overall sampling frame (not necessarily to the conceptual population). Examples include total female victims of rape and sexual assault, totals by age group, total number of rape and sexual assault events. Means and proportions can be derived as the ratios of these totals or as ratios of estimated totals to external data. Differences in rates and proportions may be addressed directly or through regression models with or without adjustment other factors.

Unbiased estimates of the variances of estimated totals can be obtained directly when applying probability sampling. Approximate variances of other statistics (means, proportions, differences, or model parameters) may be obtained through Taylor series linearization or replication methods.

The subsequent discussion of constructing survey estimates when using multiple frames will be limited to the estimation of totals.

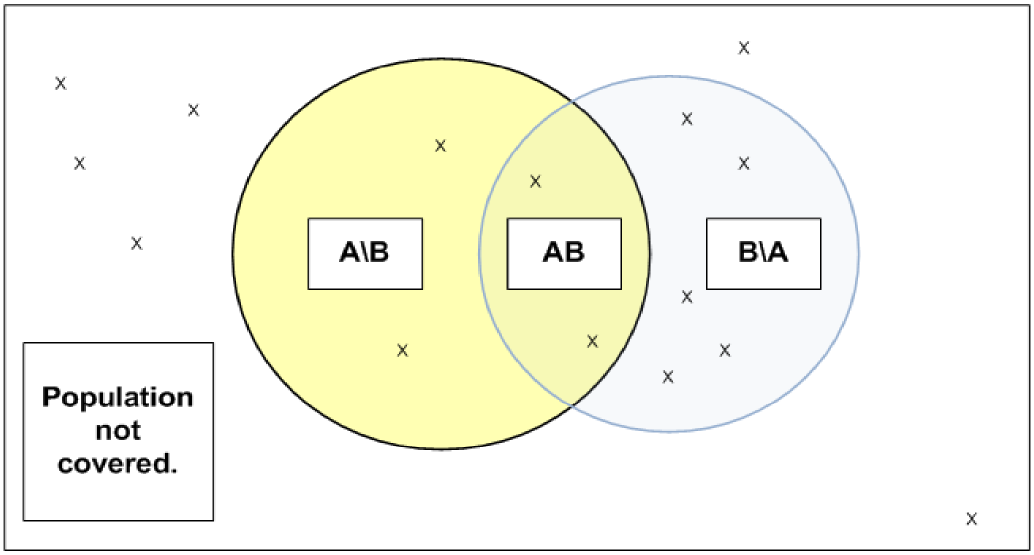
5. Overview of multiple frame methodology

For simplicity, two frames are labeled as frame A and frame B. Neither frame may cover the entire target population. Some, but not all, population members may belong to both frames. This is illustrated in Figure 1. Three subsets of the population are defined as

- Those in frame A, but not in frame B (A\B).
- Those in frame B, but not in frame A (B\A).
- Those in both frames (AB).

The whole population is contained in the outside rectangle. The X's can be viewed as population elements not contained in either frame; they represent the remaining under-coverage.

Figure 1. Population Coverage With Two Frames



A multiple frame sampling approach could be taken to improve coverage or to improve cost efficiency. Consider the coverage improvement goal and suppose a survey utilizing frame A is the primary source of data and provides good data for a large part of the target population. The coverage from frame A alone equals the ratio of the number of elements in frame A to the number of elements in the population and can be expressed as

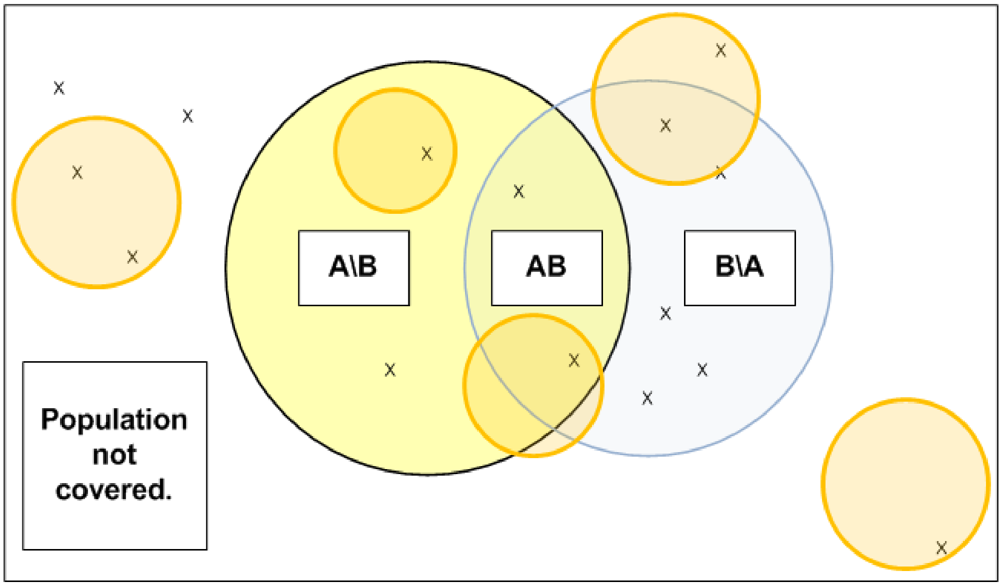
$$C_A = \frac{\sum_{i \in A}}{\sum_{i \in POP}} .$$

Adding frame B expands the coverage to all elements within either frame A or frame B or the union of the two frames ( ). The coverage achieved by using two frames can be expressed as

The second term in the partitioned sum represents the increase in coverage achieved by adding frame B.

Consider adding frames sequentially. Adding any frame that contains elements not contained in the union of the previously include frames will increase coverage, but increases in coverage are likely to diminish as the process continues. Figure 2 shows five additional frames that intersect to varying degrees with the union of frames A and B. Even those that are fully contained within the union of previous frames may be considered if they can be used to reduce the cost of achieving a specified precision level.

Figure 2. Increasing Coverage by Adding More Frames



6. A Dual Frame Approach Using Hartley's Approach

Following the dual-frame setting described previously, the total for some variable Y is limited to the population covered by the union of the two frames, A and B. Its simplest form is

$$= \quad .$$

The following discussion is based on the methods introduced by H. O. Hartley (1962). The total can be decomposed into three parts.

$$= \quad .$$

A probability sample is drawn from each frame. When drawing from frame A, we may or may not know in advance which elements are in the two parts of A: and If we do know we should stratify A into these two groups and optimize the overall design using this knowledge. In the more general case, we would not know in advance, but could determine subset membership during the observation or data collection stage. In this case, totals for the two subsets of frame A can be estimated by treating them as domains. Note that for estimating totals, one or the other of these domains might contain no sample elements resulting in an estimate of zero. An estimate of zero is not a problem theoretically; in some cases, researchers may assume that the additional coverage provided by an additional frame is very low, but they wish to claim coverage so they can obtain unbiased estimates for the larger union of frames.

From Frame A alone, the following estimates are possible:

Similarly, from Frame B alone, the following estimates are possible:

.

Note that the weights used in each estimator component depend on the frame from which the element was selected. Table 2 shows the possible estimates by domain.

Table 2. Domain Estimates Defined in a Dual Frame Example

Frame	**Population Elements in Frame A Only**	**Population Elements in Frame B Only**	**Population Elements in Both Frames (A∩B)**
Frame A			
Frame B			

The estimates from frames A and B can then be combined to produce an estimate with greater coverage.

where . The mixing parameter, λ, can be preset to minimize the expected variance.

For the estimate is based primarily on frame A, with frame B just providing a supplement of those frame B elements not included in frame A. The converse is true for . This approach is particularly applicable when the researcher knows in advance the frame membership in both frames. An example of known frame occurs in telephone number sampling. A special targeted frame may be available from specialized vendors, but cannot be assumed to have good coverage. A supplement from a list-assisted random digit dialing frame (RDD) can improve the coverage. The domains shown in Table 2 can then be treated as sampling strata. Only numbers not on the targeted telephone sampling frame would be included from the list-assisted RDD frame (e.g, Roman et al 2010 and Black et al 2011).

In the more general case, membership in frames can only be ascertained after the data are collected. For planning purposes, it helps to assume the size of each of the three components (A\B, AB, and B\A) and the level of their overlap even though the set membership of individual elements can only be known after they are observed. A few numerical examples based on specified assumptions helps to show when a multiple frame approach can both improve population coverage and maintain a specified precision level.

Three population distributions across two frames were defined for study and, in all cases, it was assumed that the union of the A and B frame covers the population adequately. Distribution 1 assumed 80 percent of the population was represented in A\B, 10 percent in AB, and 10 percent in B\A. Note that frame A alone would cover 90 percent of the population and frame B alone would cover 20 percent of the population. A dual frame approach relative to frame A only would increase coverage from 90 percent to 100 percent.

Distribution 2 assumed 50 percent of the population was represented in A\B, 45 percent in AB, and 5 percent in B\A. Under these assumptions, frame A alone would cover 95 percent of the population and frame B alone would cover 50 percent of the population. A dual frame approach relative to frame A only would increase coverage from 95 percent to 100 percent.

Distribution 3 assumed 50 percent of the population was represented in A\B, 25 percent in AB, and 25 percent in B\A. Under these assumptions, frame A alone would cover 75 percent of the population and frame B alone would cover 50 percent of the population. A dual frame approach relative to frame A only would increase coverage from 75 percent to 100 percent.

Five cost configurations were studied. It was assumed that the marginal cost of increasing the sample size for our target population using frame A was high. The high value in the simulation is represented by 1000. Frame B costs would be expected to be lower, but an equal cost option was included for completeness. Letting the frame A costs remain fixed at 1000, frame B costs were set at 1000 (1:1), 100 (10:1), 20 (50:1), 10 (100:1), and 1 (1000:1).

An approximate variance model was developed to express the variance as a function of the sample allocation to each frame and the mixing parameter, λ. By assuming a homogeneous population variance, across the three post-strata defined by A\B, AB, and B\A, this function may be expressed as

where the model variance functions associated with the sample sizes for frame A and frame B are

.

The W's represent the expected population distribution to the post-strata: A\B, AB, and B\A. The variance was constrained to be This upper limit on the variance is equivalent to requiring an effective sample size of 1000. For any value of λ, an optimum sample allocation may be expressed as:

and

The optimum mixing coefficient, λ, was approximated by testing values between zero and one at one tenth intervals. More precise options are possible, but these appear sufficient to illustrate the properties of dual frame sampling. Table 3 shows the results of the optimization process.

Under all three scenarios, satisfaction with the coverage provided by frame A would lead to cost of 1,000,000 for an effective sample size of 1000. If we assume equal unit costs for both frames (scenario 1a), the same precision could be achieved with additional coverage of 10 percent of the population for a cost increase of about 4 percent. As the relative unit costs of frame B decrease, the same level of precision can be maintained by small decreases in the frame A sample and large increases in the frame B sample. The optimum allocation to frame A remains fairly high even with the 1000:1 cost ratio of scenario 1e.

Under scenario 2a, equal unit costs, maintaining the precision level and adding coverage of an additional 5 percent of the population increases costs by over 8 percent. Since a large portion (45 percent) of the population is included in both frames, the higher probabilities of selection for elements in AB lead to a substantial unequal weighting effect. As the cost ratio increases, larger savings can be achieved by substantial increases in the sample selected from frame B.

Table 3. Approximately Optimal Allocations For Dual Frame Example for an Effective Sample Size of 1000

	Population Distribution			Marginal Cost Components		Mixing Parameter	Sample Sizes		
Scenario	A\B	AB	B\A	Frame A	Frame B	λ	Frame A	Frame B	Variable Cost
1a	0.80	0.10	0.10	1000	1000	0.6	885	155	1,039,840
1b	0.80	0.10	0.10	1000	100	0.3	775	496	824,241
1c	0.80	0.10	0.10	1000	20	0.1	744	1,178	767,316
1d	0.80	0.10	0.10	1000	10	0.1	737	1,652	753,571
1e	0.80	0.10	0.10	1000	1	0.1	726	5,145	731,153
2a	0.50	0.45	0.05	1000	1000	0.6	825	257	1,081,629
2b	0.50	0.45	0.05	1000	100	0.2	583	1,081	691,390
2c	0.50	0.45	0.05	1000	20	0.1	524	2,436	572,562
2d	0.50	0.45	0.05	1000	10	0.1	511	3,359	544,381
2e	0.50	0.45	0.05	1000	1	0.0	486	11,147	497,044
3a	0.50	0.25	0.25	1000	1000	0.5	679	413	1,091,615
3b	0.50	0.25	0.25	1000	100	0.2	471	1,091	580,102

3c	0.50	0.25	0.25	1000	20	0.1	418	2,291	463,992
3d	0.50	0.25	0.25	1000	10	0.1	406	3,146	437,539
3e	0.50	0.25	0.25	1000	1	0.0	385	9,932	394,615

Under scenario 3, frame A covers only 75 percent of the population. Under the equal cost assumption, extending coverage to the remaining 25 percent requires an increase of costs exceeding 9 percent. Frame B covers 50 percent of the population. As frame B unit costs decrease, heavier reliance on the frame B sample produces substantial savings.

For all three scenarios, the nominal sample size when frame B costs are low relative to frame A costs is quite high with the largest part coming from frame B. As a result, estimates about the aggregated population will have high design effect. This may less of a problem for domains that are found primarily from sampling frame B.

7. Other multiplicity estimators

The dual frame methodology of Hartley discussed above provides a convenient model for studying the properties of dual frame sampling. The weight applied to a sample observation under the Hartley scheme depends on the frame from which an element was selected and for elements in the intersection of the frames, AB, it also depends on the mixing factor, λ. If the same element in AB is drawn from both frames, it can contribute to both and and have a different weight depending on the frame from which it was drawn.

Two alternative methodologies and their estimators are discussed below. The first approach produces a Horvitz-Thompson-like (HT-like) estimator where the weight applied to a sample element depends on its overall selection probability. The second approach uses the concept of multiplicity to adjust sample weights to avoid double representation even though elements appear more than once in the combined frames or even within the same frame. This approach was studied extensively by Sirken (1970).

Horvitz-Thompson-like Estimators: Horvitz and Thompson's (1952) estimation theory applies a weight to each sample element that is the inverse of its overall selection probability. For the dual frame case, probabilities of selection for each element, *i*, are defined for each frame. Let be the probability of selecting element *i* from frame A, and be the probability of selecting element *i* from frame B. Applying Horvitz-Thompson theory strictly, unit *i* would be counted only once whether it was selected from just one frame or from both. Under this strict interpretation of the theory and assuming independent selection from frame A and frame B, the overall probability of selecting unit *i* would be

.

The Horvitz-Thompson estimator of a total, = , is .

Use of a generalized version of the Horvitz-Thompson estimation (Chromy 2009), would allow an element selected from both (or multiple) frames to be represented in the sample for each time selected. In place of using the overall selection probability, this approach requires computation of the expected number of times an element would be selected under the sample design, . For the dual frame example

.

Note that this value is not a probability and may exceed 1. The generalized Horvitz-Thompson estimator of = is . For the dual frame estimator this estimator can be partitioned into separate sums over each frame as

.

This approach allows the variances to be estimated for each frame and summed to get the variance of the combined total with expanded coverage.

Sirken's Multiplicity Estimators: Both the Hartley approach and the Horvitz-Thompson –like approach can be generalized to more than two frames, but the multiplicity estimators studied by Sirken make the formulation of the estimators much simpler. It helps to conceptualize the population total as

= ,

where is the number of times element *i* appears in the combined frames. For the dual frame example, it would take on values of 1 in A\B and B\A and 2 in AB. A necessary condition is that

.

The simplest choice for the is to set them all to , but any positive values between 0 and 1 may be used. In particular, for the dual frame example above, the could be set as part of the sample allocation to minimize cost while maintaining the required precision.

For the dual frame example the estimator of the population total can be expressed as

.

The concept of multiplicity can be extended to duplicates within the same frame as well as common membership in more than one frame or to network sampling where the multiplicity would be the size of an element's network.

Since Sirken's multiplicity approach can be adapted to imitate Hartley's dual and multi-frame approaches and can handle duplicates and a large number of frames fairly easily, it appears to provide the best framework for application to the design of a sampling scheme for incorporating a wide variety of sources for measuring rape and sexual assault.

8. Protection of Respondents with Supplemental Frames

While statistical methodologies may be used with multiple frames to improve estimates of totals, the heinous nature of rape and sexual assault crimes requires that special consideration must be given to any study or survey that seeks information from, or about, victims of such crimes. In particular, consideration should be given to how potential study participants will be contacted and enrolled, how they will be interviewed, how survey data will be disseminated, and how study participant identities will be protected. We discuss some particular issues related to constructing and using supplemental frames for use in surveys on rape and sexual assault.

Use of administrative data

Administrative data sources, such as police, hospitals, courts, and rape crisis centers, may have access to sufficient information to provide an alternate means of identifying victims of rape or sexual assault for the purposes of constructing a supplemental sampling frame. In such situations, a list of victims constructed from these administrative data sources could be directly compared with a list of victims constructed from another source, such as the NCVS or NISVS, in order to determine the overlap between the frames.

There are (at least) two situations where administrative data sources may be useful even when lists of victims may not be accessible; the first situation is where administrative data is used only to expand population coverage beyond that offered by the other available frames and the second is where it can be determined if sampled members from other frames are represented in the administrative data. For illustration, suppose a nationally representative sample of police stations is selected and that administrators of each station agree to report the total number of homeless females, aged 12 or older, who reported being raped or sexually assaulted in 2010. The totals reported by the police stations may be combined with NCVS totals to estimate the number of females, aged 12 or older, living in group quarters, households, or who are homeless who were raped and sexually assaulted in 2010.

Now suppose that the NCVS adds a question to a future survey to determine if respondents made a police report about being raped or sexually assaulted in a given year. Using a nationally representative sample of police stations, if the administrators of each station agreed to report the number of non-homeless females, aged 12 or older, who reported being raped or sexually assaulted in the same year as the NCVS survey, then the police provided totals may be combined with the NCVS totals, after accounting for multiplicities, in order to arrive at adjusted national estimates of the total number of females, aged 12 or older who live in group quarters or households, who were raped and sexually assaulted.

The degree to which administrative data may be used in these two situations does depend upon whether or not it is possible for the owners of the data to identify victims that meet the necessary criteria for inclusion in totals. Also, it may be easier for administrative data owners to report information on the total number of victimizations instead of total number of victims. Even with these caveats, from a practical standpoint, administrative data sources can be useful even if lists of victims are not available.

Obtaining permission to interview

While victims of sexual assault and rape may be known to police, hospitals, the courts, rape crisis centers, or other agencies, they may be unable or unwilling to provide contact information for research. Agencies may be unable to provide contact information because of State or Federal privacy laws, such as the Health Insurance Portability and Accountability Act (1996) or the Privacy Act (1974), and may otherwise be unwilling due to the sensitive nature of sexual assault and rape crimes.

Where agencies are not prohibited from providing contact information but are unwilling to directly share contact information, it may be possible to enlist the agencies to seek permission to interview. Of course, the ability of agencies to carry out such activities depends upon, among other things, whether or not the agencies have personnel trained to interview victims of rape and sexual assault and, in particular, have personnel trained to provide informed consent.

Timing of interviews relative to reported event

If the use of supplemental frames will require interviewing victims from the supplemental frames, as opposed to just noting whether or not previously interviewed victims from, say, an existing survey, are in the supplemental frames, then timing of interviews must be sensitive to when crimes occurred. Victims of rape and sexual assault are subject to psychological trauma as well as physical trauma and, as such, care should be taken when interviewing victims so that interviews do not occur too soon after an assault or rape. For the purposes of estimating rates of sexual assault and rape over a reasonable lengthy period of time, say a year, it is not necessary to interview victims immediately or in the near term after an assault occurs. For example, the NCVS uses a 6 month recall period while the NISVS uses a 12 month recall period.

Respondent protection

Studies that seek to collect information from, or about, victims of rape and sexual assault should develop procedures to ensure that study participants are provided sufficient information to provide informed consent. Furthermore, all aspects of such studies should be designed and reviewed in order to ensure that sufficient care is taken to prevent additional harm coming to study participants. An Institutional Review Board or other similar committee should provide oversight functions in order to ensure that all research is ethical, scientific, and follows appropriate regulatory requirements.

8. Conclusions

Current surveys that provide information on rape and sexual assault have low yields of victims of rape and sexual assault. If supplemental frames can be constructed from other sources of information on victims of rape and sexual assault then the precision of national estimates of totals may be improved. Sirken's multiplicity estimator provides a straight forward method for combining total estimates from multiple frames. The largest barrier to using supplemental frames to improve national estimates is lack of direct access to other sources of information about victims of rape and sexual assault.

Even when direct access to lists of victims from administrative sources is not available, administrative data sources can be useful in situations where the owners of the administrative data sources report total number of victims (or victimizations). The utility of administrative data sources to improve estimates, say derived from the NCVS and NISVS, is dramatically increased if it can be determined whether or not respondents to the surveys are represented in the administrative data sources. If particular administrative data sources are of interest, future rounds of the NCVS and NISVS could incorporate additional survey questions to determine if survey respondents are represented in the particular administrative data sources.

References

United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. National Crime Victimization Survey, 2010. ICPSR31202-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2012-06-22. doi:10.3886/ICPSR31202.v2

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.

Truman, Jennifer L., "Criminal Victimization, 2010." *National Crime Victimization Survey*. NCJ 235508, Washington, DC: United States Department of Justice, Bureau of Justice Statistics, Sep 2011.

Hartley, H. O. (1962), "Multiple Frame Surveys," *Proceedings of the Social Statistics Section, American Statistical Association*,, 203-206.

Roman, A. M., et al. (2010), "Using Lists with RDD Samples: An Examination of Bias, Cost, and Variance Estimates," *Proceedings of Survey Research Methods Section*, *American Statistical Association*.

Sirken, M.G. (1970), “Household Surveys with Multiplicity,” *Journal of the American Statistical Association*, 65, 257-266.

Horvitz, D.G., Thompson, D.J. (1952), "A generalization of sampling without replacement from a finite universe," *Journal of the American Statistical Association*, 47, 663-685.

Chromy, J.R. (2009), "Some generalizations of the Horvitz-Thompson Estimator," *Proceedings of the Survey Research Methods Section, American Statistical Association*, 217-227.

Health Insurance Portability and Accountability Act of 1996, 42 U.S.C. § 1320d-9 (1996).

Privacy Act of 1974, 5 U.S.C. § 552a (1974)