A Multiple Frame Approach to Sampling the Homeless and Transient Population

Ronaldo Iachan and Michael L. Dennis[1]

Abstract: This paper describes the design of the homeless study component of the Washington, DC, Metropolitan Area Drug Study (DC*MADS). The study adopted a multiple frame approach to cover the homeless population. The frames included shelter and soup kitchen frames as well as street and encampment frames. The design also included temporal stratification with the selection of independent monthly samples of institutions, street blocks and sample days, and the random assignment of sample units to days. The paper discusses issues related to sampling homeless people, combining data across multiple overlapping frames, and the relative coverage of the different frames. It then examines the implications of these findings for future studies.

Key words: Rare mobile population; street sample; shelter sample; temporal stratification.

1. Introduction

Over the last decade, growing attention has been devoted to the homeless problem which has been steadily increasing in scale as it has changed in character in the United States and other countries. Recent studies have investigated the numbers and characteristics

of homeless individuals in different areas of the United States. This paper explores design issues relevant to homeless surveys. These issues are discussed and illustrated with a study of homeless people in the Washington, DC, metropolitan area currently conducted by the Research Triangle Institute as part of the National Institute on Drug Abuse's, Washington, DC, Metropolitan Area Drug Survey (DC*MADS) study. The Washington, DC homeless survey includes shelter, street and service location components.

Two issues that should be considered in the design of surveys of homeless street people pertain to the need for stratification both by geographic location and by time. Spatial stratification is necessary to help locate eligible members of this population, which is both rare and mobile. Temporal stratification allows both the computation of estimates of prevalence and incidence,

and of seasonal variations and trends. A third issue relates to the use of multiple sampling frames, to capture populations components in such disparate settings as shelters, service locations, and streets.

Shelter surveys only capture a small portion of the homeless population and do not properly represent subgroups of potential interest (Dennis and Iachan 1991). For example, Davidson (1991) found that the rates of substance abuse, mental illness and mental retardation among 313 people served by nine shelters on one day in July were significantly different from the rates among the 632 who were on the shelters "do not admit" list on the same day (i.e., records of people refused admission for various reasons related to previous or current behavior). Several studies have attempted to address this bias by supplementing shelter surveys with samples of people drawn from other locations. Dennis (1991) categorized 14 homeless studies into three groups. Studies in the first category use only samples of service system locations (e.g., shelters, soup kitchens, day programs) because they are cheaper and cover most of the population (e.g., Breakey et al. 1989; Burt and Cohen 1989). Those in the second group consider probability samples of shelter and street locations to reduce the potential for bias due to undercoverage and limitations of service systems (e.g., Rossi, Fisher, and Willis 1986). Studies following the third, compromise approach, focus on service system samples but also include either purposive or partial samples of high-density street locations (e.g., Vernez, Burnam, McGlynn, and Mittman 1988; Farr, Kogel, and Burnam 1986; Ringwalt and Iachan 1990). Only one study, the DC*MADS Homeless Study reviewed in the next section, attempts a comparison of all three approaches.

2. An Illustrative Homeless Survey

For almost two decades, NIDA has relied on a series of household and hospital surveys to monitor substance abuse in America. While this strategy has been useful as a general barometer of drug use, concern has increased that it underrepresents several subpopulations that are more likely to be adversely affected by substance abuse such as school dropouts, adult and juvenile criminal offenders, the institutionalized, drug abuse treatment clients, pregnant drug abusers and, most notably, the homeless population.

NIDA has contracted with the Research Triangle Institute and three firms to conduct a series of 16 comprehensive studies under the umbrella of a single research study program, the DC*MADS. This effort is an attempt to collect data about drug abuse from all of these subpopulations and the household population during the same year in one metropolitan area. To the extent that it is successful, DC*MADS will be used as a model to collect similar data in other metropolitan areas.

The DC*MADS homeless study component (NIDA 1993) examines the prevalence, and consequences of drug use in the homeless population. The study allows comparisons of several definitions of homelessness. At one extreme, it includes all people encountered in encampments, shelters and service locations, like the homeless enumeration study conducted by the U.S. Census on March 20, 1990. At the other extreme, the target population can be confined to literally homeless persons who at a given night may be found either in a shelter or on the streets. It may be worth pointing out that neither definition necessarily includes all doubled-up persons in households (i.e., persons staying with relatives or

friends), institutionalized persons or persons considered at-risk for homelessness.

It is also important to consider the time dimension associated with the homeless definition. Expanding this time window can dramatically increase the estimated population size and change its characteristics. Santiago, Bachrach, Berren, and Hannah (1988), for instance, found that changing their definition from "currently homeless" to "homeless in the last three months" increased from 106 to 159 (50%) the number of people identified as homeless in a sample of psychiatric hospital patients. Additional discussion of definitional issues may be found in Hopper (1990, 1991).

The DC*MADS study targeted people who were either literally homeless or at imminent risk of moving into homelessness. An eligible person met one of the following conditions on the sampled day:

- someone who stayed overnight in an emergency shelter for homeless people, runaways, or neglected or abused women;
- someone who stayed overnight in a house, apartment, or room paid for with municipal emergency housing funds;
- someone who stayed overnight in a nondomicile, such as a vacant building, public or commercial facility, city park or car, or on the street;
- someone whose regular place to stay was a nondomicile regardless of where he/she stayed the prior night (e.g., people who traded sexual favors for shelter or spent one night in a hotel or hospital); or
- someone who was using a soup kitchen or emergency food bank for the homeless population.

Note that our definition of homelessness includes many people who are precariously

housed or living in nontraditional arrangements (e.g., trading sex for shelter). Unlike some other studies, we chose not to include day programs as a sampling site because in the Washington, DC metropolitan area they were typically operated by either a shelter or soup kitchen, or primarily served other populations (e.g., juveniles, the elderly).

A screener was administered to potential respondents in the street and encampment settings to exclude those who were not literally homeless. This screener was used at the soup kitchen sites to allow a subgroup of literally homeless people to be identified, but it was not used to exclude people. It was not necessary to use the screener among shelter users, who were automatically defined as literally homeless.

People who were cognitively impaired and could not complete the interview were also excluded. Impairment was defined by evidence of extreme intoxification (e.g., passing out) or scoring more than nine on the Short Blessed Exam (Katzman, Brown, Fuld, Peck, Schecter, and Schimmel 1983). Screener questions are given in the Appendix.

3. Sample Design

The sample design involved surveying people sampled over both time and space to develop estimates and characteristics of the homeless and transient population on an "average" day between February and June 1991. Four temporal samples of 16 days each were selected in the months of February, March, April, and June 1991, at a rate of four days per week. Spatially, there were one to two samples (with replacements) from each of the four sample frames: two samples of shelters, one sample of soup kitchen meals, one sample of encampment clusters, and two samples of street census

tracts and blocks. Table 1 summarizes the sample design, actual sample, and response rates for the four study components.

The months and number of days vary among the sampling frames because the study design was changed at the end of April 1991 to address several problems with the street component. These problems included higher than expected physical risks to interviewers (including armed robbery), lower than expected numbers of completed interviews, and higher than expected overlap between the street and alternative sampling frames. The initial design used random samples of shelter and street people similar to the Robert Wood Johnson Chicago study on homelessness (Rossi et al. 1986). In the redesign, which was implemented in June 1991, the shelter sample was maintained, but the street

sample was replaced with samples from soup kitchens and encampments.

The temporal sample was selected as a stratified random sample of 16 days for each of the four (four-week) months: February, March, April, and June. Temporal strata were weeks with four days randomly sampled within each week. First-stage, spatial sample units (shelters, soup kitchen meals, encampment clusters, census blocks) were randomly assigned to sampled days to avoid temporal bias in the data collection. The estimate for each sampled day is an unbiased estimator of the homeless population total, so that the mean over the sampled days is an unbiased estimate of the average daily number of homeless people (during a given month or the entire data collection period). This approach also eliminates the need to estimate multiplicity across sampled

Table 1. Sampling design summary for the street and shelter survey components

Sampling stage/unit	Sampling method	Size (rate)
1. *Shelter sample*		
1a. Days	Stratified random sampling	64 days (4 per week)
1b. Shelters	Probability proportional to size (bed capacity)	94 shelters (1–2 per day)
1c. Shelter clients	Systematic random sampling	484 interviews (5–6 per shelter)
2. *Street sample*		
2a. Days	Stratified random sampling	48 days (4 per week)
2b. Blocks	Stratified random sampling	432 blocks (9 per day)
2c. Street homeless persons	All eligible individuals found in sample blocks/days	54 interviews (.125 per block)
3. *Encampment sample*		
3a. Days	Stratified random sampling	16 days (4 per week)
3b. Encampments	Simple random sampling	16 encampments (1 per day)
3c. Encampment homeless persons	All eligible individuals found in sample encampment/days	146 interviews (9–10 per encampment)
4. *Soup kitchen sample*		
4a. Days	Stratified random sampling	16 days (4 per week)
4b. Soup kitchens	Probability proportional to size (meal capacity)	32 kitchens (2 per day)
4c. Soup kitchen clients	Systematic random sampling	200 interviews (6–7 per kitchen)

Source: Adapted from NIDA (1991).

days (Dennis, Iachan, Thornberry, and Bray 1991). An overview of each of the four main components is provided below. More detail on building and evaluating the sampling frames follows this summary.

The shelter sample was selected in two stages. At the first stage, a stratified sample of shelters was selected from the seasonal frame. Both winter and spring samples were stratified by size, the size measure being shelter capacity. At the second stage, residents of each selected shelter in the sampled night were selected with equal probabilities from an intake or resident roster using systematic random sampling. The shelter sampling rates were specified using expected occupancy data collected from each sampled shelter prior to the assigned data collection night. The sampled shelters in each month were randomly assigned to 1 of 16 randomly sampled nights.

The soup kitchen sample was selected in two stages. At the first stage, site meals (e.g., breakfast, lunch, dinner) were selected with probabilities proportional to size, the size measure being the expected number of persons (meals) served for a single sitting of a particular meal at a single site. Note that many soup kitchens had multiple sittings of each meal or operated programs at multiple sites. An equal number of persons was selected from each sampled meal unit so that the final sample would be approximately self-weighting. The sampled soup kitchens were randomly assigned to the 16 sampled nights in June 1991.

Table 2 summarizes the sample sizes for the shelter and soup kitchen samples.

Table 2. Institutional sample summary, adjustment, and response rates

Sample status	Shelters			Soup kitchens/meals[1]	Institution total
	Winter	Spring	Total		
Sampling frame	290	290	290	105	395
Initial	63	60	123	32	155
Ineligible	6	12	18	5	23
Refused	10	11[3]	21[3]	1	22[3]
Completed	47	37[2]	84[2]	26	110[2]
Replacement	13	1	14	5	19
Ineligible	2	0	2	0	2
Refused	4	0	4	0	4
Completed	7	1	8	5[2]	13[2]
Total sampled	76	61	137	37	174
Ineligible	8	12	20	5	25
Refused	14	11[3]	25[3]	1	26[3]
Completed	54	38[2]	92[2]	31[2]	123[2]
Institutional response rate[4]	79.4%	77.6%	78.6%[3]	96.9%	82.6%

[1]The sitting of a particular meal at a given location was the sampling unit; thus, two sittings of breakfast at the same program were treated as two units for sampling purposes.
[2]Includes one institution that agreed to participate but had no clients on the sampled day.
[3]Includes three shelters that refused participation in the winter and were not recontacted when they were resampled.
[4](Total completed)/(Total completed + Total refusal).
Source: 1991 NIDA DC*MADS Homeless and Transient Population Study.

Encampments were identified by the same local experts who provided the ratings for the street sample (discussed below) and who helped update the shelter and soup kitchen frames. Encampments, defined in terms of as many contiguous census blocks as necessary (e.g., along Pennsylvania Avenue between 10th and 15th Streets), were visually verified by means of a drive-by. Although encampment clusters typically included several urban census blocks, the largest cluster was smaller than several of the individual enumeration districts and rural blocks in the street sample.

The two seasonal street samples were selected in three stages; sample tracts and blocks were selected with stratified random sampling in the first two stages. Tracts and blocks (in the sampled tracts) were classified as high, medium, or low according to expert ratings (discussed below). The experts for the tract ratings were individuals with primary responsibility for providing homeless services in each municipality. The experts for the block ratings were nominated by officials who were familiar with the specific tracts sampled in the municipality (e.g., shelter operators, outreach workers, Health Care for the Homeless staff). All individuals found in the selected blocks in the sampled periods were included in the final sample and screened. Note that the two block samples, selected for the winter and for the spring data collection, were selected from a single sample of tracts. The sampled blocks were randomly assigned to one of the two months in the season, grouped into 16 geographic clusters and then randomly assigned as a cluster to 1 of the 16 sampled days in the month.

4. Sampling Over Time

Time-related problems that must be addressed in sampling homeless people include:

- seasonal changes, e.g., due to weather,
- changes in the service systems,
- population movement across sampling frames, and
- problems associated with using a currently-homeless definition.

Seasonality affects the number and distribution of homeless people in many ways. In most areas, winter means higher utility bills that force some people out of their homes. On the other hand, warmer spring weather makes sleeping outside a more viable option. The seasonality of the data collection period underlies the relative allocation of the total samples to shelter and street sites. Because more people seek shelter in cold weather, more observations are required from shelters than from the streets in the winter, and conversely in the spring.

We selected independent seasonal samples stratified by month, and randomly assigned shelters and blocks to the sampled nights. The selection of monthly samples prevents clusters of days at the beginning or end of the season. It also minimizes the chance of visiting all of the selected shelters or blocks in one municipality in the same month. More importantly, it permits the computation of monthly and seasonal estimates and trends (Iachan 1989).

Most of the service systems in the Washington, DC area change their level of services around April 1 of each year. The two independent seasonal samples are designed to capture the April 1 change in the service systems. To avoid having too many days clustered around the beginning or end of the month, when entitlement checks and paychecks often arrive in the mail, the temporal sample is stratified by week. Potential biases are further reduced by randomly assigning shelters and blocks to the sampled nights. Even stronger arguments for temporal stratification can be

made in other cities or places with more pronounced seasonality and migration patterns.

In many Washington, DC shelters, people are entering shelter buildings from 6:00 p.m. until 6:00 a.m. However, even before everyone is in for the night, many start leaving (up to 50% of shelter clients may have left by 4:00 a.m. to start walking over to a soup kitchen). Thus, there is no one single time in which the entire shelter population for a given night can be captured. It also means that on a given night, the same person may be in different shelter and street frame units. The NIDA survey addressed the first problem by taking a systematic sample of people as they enter the shelters throughout the night. The same sample nights were used for the street and shelter samples to minimize the overlap between the two components. The chances are negligible that a person can be found in a shelter and then in the street between 4 and 5 a.m. that same night. The street data collection takes place in a period of relatively low mobility (4:00 to 5:30 a.m.). Finally, the respondents were asked whether they have ever been interviewed before, and other questions that permit multiplicity adjustments to be made (see Section 6).

Determining the overlap between multiple frames is a common problem in designing a sample. When an overlap cannot be defined away, it is necessary to measure it in developing a population estimate. In homeless studies, this has been done by asking people about their sleeping quarters and service utilization in the last 7–30 days (e.g., Burt and Cohen 1989; Farr et al. 1986). A common but more dubious practice is to inflate or extrapolate this number to the last 12 months or a lifetime. The problem with the latter technique is that the same individual often becomes homeless at several points in a year. These episodically homeless people bias the resulting adjustments and produce annual estimates of unique episodes, not unique individuals.

5. Street Sampling

Ideally a street sample should identify most of the non-domiciled people; however, it is difficult even for knowledgeable people to predict where homeless street people will be sleeping on a given night. The street population is rare, mobile, and elusive. It is difficult to locate people who are actively hiding to avoid both victimization and being run off by authorities. Unfortunately provider and advocate estimates of the number of street people are also unreliable and vary by ten-fold or more (Farr et al. 1986; Rossi 1989; Vernez et al. 1988).

The NIDA street survey was based on a two-stage sample of census blocks. Sample blocks were selected in two stages; first-stage units were census tracts. At both stages, the sample was stratified by the likelihood of finding a homeless person in the area during the predawn or early morning hours. These hours were chosen to minimize the amount of screening needed to identify eligible individuals and to select a time when they would be least mobile. Stratification information sources included local service providers and homeless people.

The first-stage sampling frame was stratified into three categories according to the likely concentration of homeless individuals. Tracts (at the first stage) and blocks (at the second stage) were classified by local experts as low, medium or high according to the likelihood of finding any homeless individuals and the expected density of homeless persons in the area. To elicit these ratings, personal interviews were conducted with the local experts (aided by local maps). For first-stage stratification, raters of census tracts were individuals with primary

responsibility for providing services to the homeless in each municipality. For second-stage stratification, block raters were nominated by these officials or other local officials familiar with the selected tracts. These raters included shelter operators, outreach workers, and staff of Health Care for the Homeless.

We used two types of experts to identify areas where homeless people would most likely be found. First, municipal experts were asked to identify on a map the census tracts in their municipality with high concentrations of homeless street people. These municipal experts were asked to keep in mind such areas as parks, 24-hour public facilities, vacant buildings, transportation depots, and places near soup kitchens or shelters where homeless people tend to stay. The experts rated the census tracts using the following scale:

Tract Scale

2. High: One or more homeless people can be found somewhere in the tract six or more nights per week.
1. Medium: One or more homeless people can be found somewhere in the tract one to five nights per week.
0. Low: Remaining tracts.

A random sample of census tracts was selected using these ratings; tracts rated high and medium had a greater selection probability.

In the second stage, the municipal experts were asked to provide the names of people who were knowledgeable about the homeless population on a block level in the selected tract. These "community" experts included outreach workers, shelter operators, police officers, and other service providers for homeless people. The community experts rated each of the individual census blocks within the census tracts for concentrations of homeless people who could be found between 4:00 and 5:30 a.m. using a scale similar to the tract scale:

Block Scale

2. High: One or more homeless people can be found somewhere in the block six or more nights per week.
1. Medium: One or more homeless people can be found somewhere in the block one to five nights per week.
0. Low: Remaining blocks.

The community experts were reminded to think of such areas as parks, 24-hour public facilities, vacant buildings, transportation depots, and places near soup kitchens or shelters. To facilitate data collection, notes were also recorded about specific locations (e.g., behind a store, in a vacant building) where they knew people were staying.

After the first-stage stratum counts were tabulated, we developed a sample allocation that oversamples high tracts and selects medium tracts with higher sampling rates than low tracts. Sample tracts were selected with equal probabilities within each first-stage stratum. At the second stage, we adopted a similar allocation that oversamples high blocks. Sample blocks were selected with equal probabilities within each second-stage stratum. (Note that nine cells result from different combinations of block and tract ratings, and that a different sampling rate was used within each cell.) All homeless individuals identified in a sample block during the data collection period were counted, screened and interviewed.

Table 3 presents the actual number of interviews conducted in the sample blocks in each stratum. This table provides a graphic illustration of the effectiveness of the stratification. Even though a vast majority of the tracts and blocks were rated low, most interviews were conducted in the high and medium blocks (and tracts).

Table 3. Street interviews in block and tract strata

Block rating (Block sample/universe)[1]	Tract rating (Tract sample/Tract universe)			
	High (32/111)	Medium (16/117)	Low (16/602)	Total (64/830)
High (175/176)	8	47	0	55
Medium (103/146)	2	1	6	9
Low (160/1720)	0	0	0	0
Total blocks (432/2042)	10	48	6	64

[1]The winter and spring samples used the same census tract sample, but independent block samples with replacement; these are the combined block counts.

Source: 1991 NIDA DC*MADS Homeless and Transient Population Study.

Despite the improved efficiency of using expert ratings, the street survey cost more than twice as much as the rest of all the other data collection and created numerous safety problems. The June sample tested a different approach in which the local experts directly identify the places (encampments) where the experts thought homeless people could be found on virtually every night. Although local experts had been willing to stratify census blocks for the street sample, they were typically confident of only a handful of sites. Virtually all of the street interviews occurred in what were expected to be high or medium density sites, but other medium and high sites contained little or no evidence that homeless people had been there in quite some time.

Encampments include such places as parks, vacant buildings, under bridges, forests, and along rivers. Eligible areas were identified as having concentrations of homeless people spending the night and verified by research staff in drive-bys as currently being used by one or more homeless people on at least one of two nights. Unlike the earlier street survey sites, an encampment could contain more than one census block and often ran along a street or surrounded an area such as a park or a shipyard.

To identify the encampments, it was necessary to review the list of likely sites with local officials, shelter providers, outreach workers, staff of mobile soup kitchens, and field supervisors. Every site that was nominated was then verified with two separate drive-bys and any site that had two or more people who appeared homeless on either drive-by was kept. Because many sites were highly clustered (e.g., several blocks along a single street), they were combined into a single encampment cluster for data collection purposes.

The encampment approach was successful in that all clusters generated interviews for a total of twice as many interviews as the street sample. Unfortunately the encampment sample appears to represent only about 8% of the total street population and was not a representative cross-section of the street population in terms of demographics, drug use, alcohol use, or mental illness (NIDA 1993).

6. Multiplicity and Coverage Issues

Although randomly assigning sampling units to sampled days significantly reduces the potential for within-frame multiplicity (i.e., having more than one chance of being selected), the use of multiple frames raises the problem of multiplicity across frames on a given day. Next to differences in definition, this has probably been one of the

major sources of variation in the estimated size of the homeless population.

NIDA's Homeless and Transient Population Survey addressed this problem by asking respondents about their actual or expected use of shelters, soup kitchens, and being on the street during the 24-hour sampled day. Note that a sampled day started at 6:00 p.m., so shelter interviews occurred at the beginning of the day and were followed by street or encampment interviews in the early morning and soup kitchen interviews during the following day. Based on their responses, each person was classified in terms of being or not being in each of three population domains: shelter, soup kitchen, and street (of which encampments are a subset). To the extent that a person was in more than one frame, the probability of his/her selection was increased and a multiplicity adjustment was made to develop the correct analysis weight. More details on the weighting and estimation procedures, as well as additional survey findings, will be given in a forthcoming article.

Total Population = (100%)
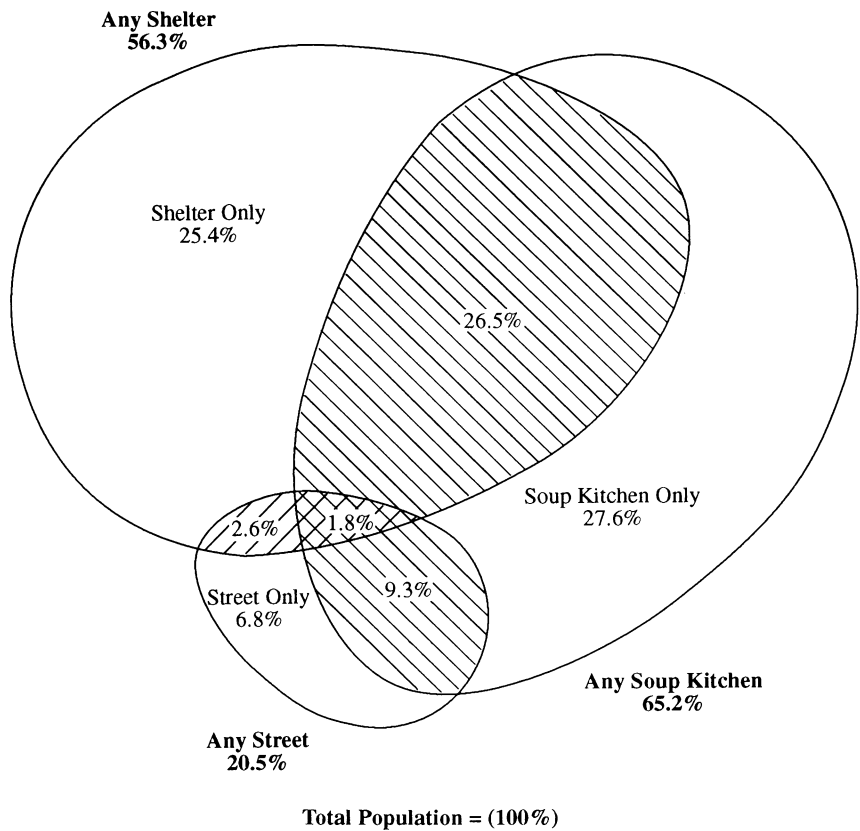
Fig. 1. Size and 24-hour overlap of the three subpopulations of homeless people in the DC MSA

Note: The total size of the homeless and transient population on an average day in February through June 1991. Failure to take even the 24-hour overlap into account would have increased the estimate by 42% and partially explains why local estimates vary so much among researchers, practitioners, and advocates of this population.

Source: NIDA 1991 DC*MADS Homeless and Transient Population Study.

Figure 1 shows the relative size of the seven population segments formed by crossing the three domains. As can be seen in Figure 1 there is considerable movement of homeless people among shelters (56%), soup kitchens (65%), and street locations (20%). Failure to adjust for double counting helps explain much of the variability in prior local estimates as it would have increased the population estimate by 42%, substantively changing from an estimate of unique individuals to an estimate of person-contacts. Keep in mind that this is the effect of multiplicity in a single day. Over longer periods of time, we would have an even greater overlap in terms of unique people and the additional problem of individuals being repeatedly sampled.

Also evident in Figure 1 is the additional coverage provided by soup kitchens (SK) for our broadly defined population. Because the population includes at-risk persons, the SK-only segment accounts for a substantial portion (27.6%) of the interviews.

From a programmatic or policy perspective, the size of the total population is often of far less interest than the size of its subgroups. Because homeless people have many different needs and reasons for being homeless, it is critical to examine the extent to which the various sampling frames cover each of the major subgroups. Note that because of the overlap in the frame, adding new frames must be evaluated by the marginal increase in population or subpopulation coverage that they offer. We examine this question in the order most researchers consider when frames are successfully added (from relative low-cost to high-cost): shelters, soup kitchens, encampments, and a probability street sample.
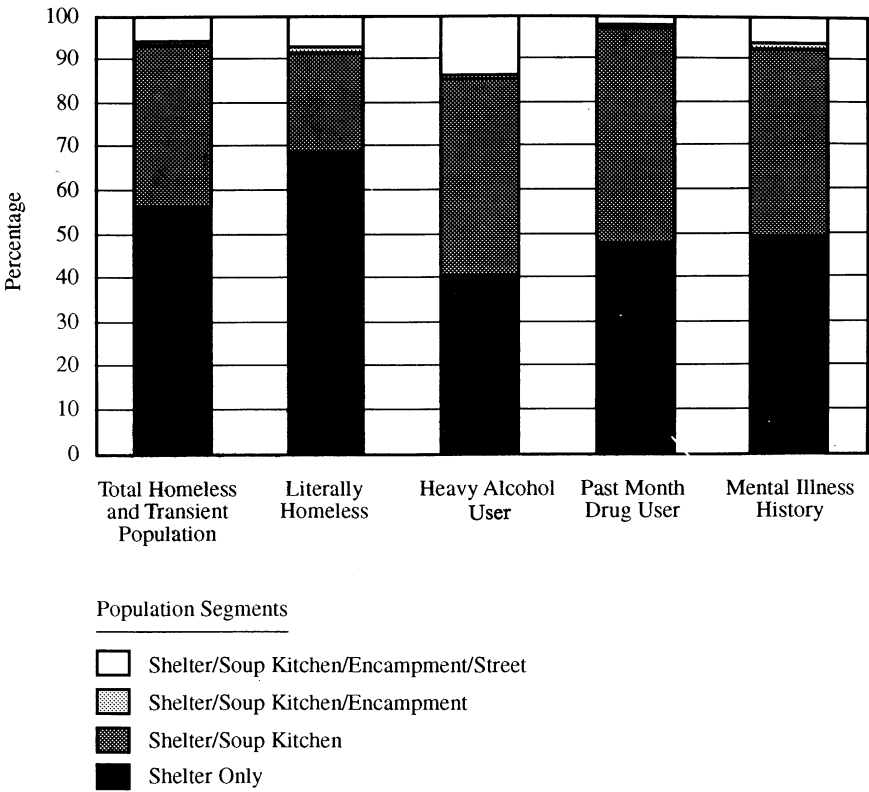
Figure 2 shows the percentages of the population or subpopulation covered by a shelter frame, a shelter plus a soup kitchen frame, those two plus an encampment frame, and those three plus a street frame. Note that the classification here is based on the population segment, not on where the actual interview was conducted. Thus, a person who reported spending the night in a shelter but was interviewed in the street would be considered "theoretically" covered by the shelters. The first two columns provide estimates for the total homeless and transient targeted population and for the subpopulations of people who were literally homeless. The next set of columns looks at the coverage of heavy alcohol users (five or more drinks per day on a weekly basis), current illicit drug users (use of marijuana, cocaine, inhalants, hallucinogens, heroin, or unprescribed psychotherapeutic drugs during the past month), and people with histories of being treated for mental illness (inpatient or outpatient).

The population coverage by frame indicates that shelters alone did not cover the total population (56%) on an average day. The addition of soup kitchens brought the coverage up to 93%. The addition of encampments added about 1% so that all three sites combined covered 94% of the total population. Three-site coverage was lower for selected groups such as youths (85%), heavy alcohol users (86%), and the unemployed (90%), however, while higher for current drug users (98%), veterans (97%), and families (95%). The results suggest that the contribution of street area samples will be typically in the 5% to 15% range and vary with the target population in question.

7. Discussion

The design of the DC*MADS homeless study street component incorporated the knowledge gained during Rossi's (1989) Chicago study and the Census enumeration (S-night). Still, the state-of-the-art design

Fig. 2.   Cumulative coverage of the homeless population, heavy alcohol users, current drug users, and people with mental illness histories

presented several problems and opportunities for further design improvements.

For the design of nighttime surveys, it may be helpful to define two primary subgroups of the street homeless population. The first group consists of those individuals who may be found clustered in encampments and who often tend to seek safety in numbers. The second group includes isolated individuals who are either wandering in drug- or mental-illness-induced stupor or are hiding for safety or privacy reasons. Locating, listing, and sampling the second group is much more difficult and expensive than the first group.

The DC*MADS survey was designed to capture individuals in both of these groups but was only partially successful in covering (or uncovering) members of the second group (NIDA 1991). This partial coverage occurred despite intensive efforts ranging from going into places of difficult or danger-

ous access (e.g., abandoned buildings and crack houses) to screening and interviewing in the hours of presumed lowest mobility. In fact, the interviewers were instructed to wait for any person found sleeping in the street to wake up. Nevertheless, a majority of the street people screened and interviewed were found in movement. Another finding of relevance is that a great majority of the eligible persons interviewed were regular service users (e.g., soup kitchens), a finding that reinforces the notion of sampling daytime service locations.

The service location sampling approach was included in the DC*MADS sample for June 1991. For this purpose, we constructed a comprehensive frame of service programs further subdivided by sites and meals. Sampling units were meal-sites (e.g., breakfast at a particular site). This study component adopted less strict eligibility criteria (i.e., broader definitions for the different degrees of homelessness), and throws further light on the overlap between the various homeless subpopulations.

Based on this review and our experience in DC*MADS, there are several alternative street sampling designs that hold some promise for further addressing the cost and precision issues related to homeless population surveys. Table 4 compares five potential strategies for sampling street homeless individuals. The relative advantages of each strategy are presented along two basic dimensions: cost and coverage. Other factors to consider include whether the strategy yields a probability sample of areas and homeless people in these areas.

The sampling frame for the first strategy consists of a list of known clusters of street persons that may be verified by field staff. For the second strategy, the frame is restricted to areas with high density ratings provided by expert judgment. To the extent that such judgments are considered sufficient to exclude an area, it will produce a partial probability sample (i.e., a probability sample of the targeted areas). Where it has been used (e.g., Vernez et al. 1988), such expert judgments have been typically verified through drive-by or "windshield" observations.

The third strategy calls for a stratified sample of areas that are then listed to exclude areas unlikely to contain homeless people. Such a procedure is analogous to that used in household surveys and would thus incorporate both expert judgment and direct observation. The fourth option is an adaptive cluster sampling method analogous to the Waksberg-Mitofsky variation of random digit dialing (RDD). The idea, as in the RDD variant, is to reduce the number of screenings

Table 4. Potential strategies for street sampling

Strategy	Cost	Coverage	Geographic probability sample	Applications
Listing/sampling encampments	Low	No loners	No	Rossi's supplement DC*MADS supplement
Sampling high-density areas	Moderate	No low-blocks	Partial	Vernez et al. (1988)
Sampling and listing	High	Very good	Yes	—
Geographic analog of Waksberg-Mitofsky	High	Fair	Yes	—
Stratified random sample	High	Good	Yes	Rossi et al. (1986) NIDA (1991) Hamilton et al. (1987)

Appendix I
Screening Questionnaire

```
┌─────────────────────────────────┐
│ FOR OFFICE USE ONLY             │
│                                 │
│ Quest. ID #   ┌──┬──┬──┬──┐      │
│               └──┴──┴──┴──┘      │
└─────────────────────────────────┘
```

DC*MADS
HOMELESS AND TRANSIENT POPULATION STUDY

SCREENING QUESTIONNAIRE

STREET SAMPLED RESPONDENT

TRACT ID#   __ __ __ __ __ __   DATE: __ __-__ __-__ __

BLOCK ID#   __ __ __ __   M M D D Y Y

INTERV. ID# __ __ __ __ __   START

TEMP. ID#   __ __ __ __   TIME: __ __ : __ __ AM/PM

                                               H H   M M

STREET SCREENER INTRODUCTION

> Hello, my name is _____, and I'm from the Research Triangle Institute. I'm working on a study in the D.C. metropolitan area for the U.S. Department of Health and Human Services. I would like to ask you some questions about housing, which will only take 1 or 2 minutes. The answers to these questions will be kept strictly confidential and no names will ever be connected with the survey.

SS-1.   First, do you have some place here in the D.C. metropolitan area that you consider to be your home or the place where you sleep regularly?

                              YES....................1 → [GO TO SS-2]

                              NO......................2

SS-1a.   Do you have some place in a different city, county, or state that you consider to be your home or where you sleep regularly?

                              YES....................1

                              NO......................2 → [GO TO MAIN Q]

SS-2. Is that a house, an apartment, a room, a shelter, a car, or a spot in some public place such as a park bench or bus station? **(PROBE UNTIL YOU GET AN ANSWER THAT INDICATES WHETHER R HAS REGULAR HOUSING OR NOT, RECORD LOCATION)**

<u>REGULAR HOUSING:</u>

A house	01
An apartment	02
A room, paid for by R	03
A boat	04
Some other form of regular housing	05
(SPECIFY) _____	

<u>OTHER ARRANGEMENTS:</u>

House, apartment or room paid for with municipal emergency housing funds	06
General shelter	07
Halfway house/transitional housing	08
Juvenile/runaway shelter	09
Domestic violence shelter	10
Public campground	11
Car or truck	12
Public facility	13
Transportation depot	14
Vacant building	15
Tunnel/sewer	16
Underpass/bridge	17
Under porch/building structure	18
Sidewalk	19
City park	20
River front	21
Forest/fields	22
Other nondomicile **(SPECIFY BELOW)**	80

→ **[GO TO MAIN Q.]**

_____

SS-3.   Where did you stay last night?

<u>REGULAR HOUSING:</u>
A house...01
An apartment..02
A room, paid for by R..03
A boat..04
Some other form of regular housing.............................05
(SPECIFY) _____

<u>OTHER ARRANGEMENTS:</u>

House, apartment or room paid for with
   municipal emergency housing funds.........................06
General shelter...07
Halfway house/transitional housing............................08
Juvenile/runaway shelter...09
Domestic violence shelter...10
Public campground...11
Car or truck...12
Public facility...13   → [GO TO MAIN Q.]
Transportation depot..14
Vacant building...15
Tunnel/sewer..16
Underpass/bridge...17
Under porch/building structure...................................18
Sidewalk...19
City park...20
River front..21
Forest/fields...22
Other nondomicile (SPECIFY BELOW)....................80

SS-4.　Who does the place you stayed at last night belong to? By belong, I mean who pays the rent or mortgage or owns it?

        Self...01 ┐ → **[SKIP TO BOX A]**
        Spouse...02 ┘

        Parent..03
        Other relative...04
        Sexual partner..05
        Friend..06
        Someone else..07
        (SPECIFY) _____

    SS-4a.　Do you have an arrangement with your (parent/relative/partner/friends/this person) to sleep in their place on a regular basis?

                YES....................1

                NO......................2 → **[GO TO MAIN Q.]**

BOX A:

Those are all of the questions I have for you. Thank you very much for your time and participation. Let me assure you again that all of the information you have given will be kept confidential. Good-bye.

INTERVIEWER: COMPLETE INTERVIEWER OBSERVATION QUESTIONNAIRE.

---

needed to find eligible population members. This reduction is achieved by following up on successful screenings in a given cluster. The fifth option is a one- or two-stage stratified random sample that incorporates expert information on the probability of identifying homeless people. Of the five listed strategies, only the last three (which are also the most expensive) assign non-zero probabilities of selection to each area. Of these, only the last one has been fully implemented.

8.　References

Breakey, W.R., Fischer, P.J., Kramer, M., Nestudt, G., Romanoski, A.J., Ross, A., Royall, R.M., and Stine, O.C. (1989). Health and Mental Health Problems of Homeless Men and Women in Baltimore. Journal of the American Medical Association, 262, 1352–1357.

Burt, M.R., and Cohen, B.E. (1989). America's Homeless: Numbers, Characteristics, and Programs that Serve Them. Urban Institute Report 89-3. Washington, DC: Urban Institute.

Davidson, D. (1991). A Snap Shot Survey of Hard to Serve Homeless Clients in Northern Virginia. Fairfax, VA: Northern Virginia Coalition for the Homeless.

Dennis, M.L., (1991). Changing the Conventional Rules: Surveying Homeless People in Nonconventional Locations. Housing Policy Debate, 2, 1–32.

Dennis, M.L., and Iachan, R. (1991). Sampling Issues in Estimating the Extent of Alcohol, Drug Abuse and Mental Illness

Problems. In C.M. Taeuber (Ed.), Enumerating Homeless Persons: Methods and Data Needs. Washington, DC.: Bureau of the Census, U.S. Department of Commerce, 188–191.

Dennis, M.L., Iachan, R., Thornberry, J.S., and Bray, R.M. (1991). The RTI Method: Sampling Over Time. In C.M. Taeuber (Ed.), Enumerating Homeless Persons: Methods and Data Needs. Washington, DC.: Bureau of the Census, U.S. Department of Commerce, 167–170.

Farr, R.K., Kogel, P., and Burnam, A. (1986). A Study of Homelessness and Mental Illness in the Skid Row Area of Los Angeles. NIMH Grant No. iHR MH36809-01. Los Angeles, CA: Department of Mental Health.

Hamilton, Rabinovitz, and Alschuler, Inc. (1987). The Changing Face of Misery: Los Angeles' Skid Row Area in Transition. Four volume report to the Community Redevelopment Agency of Los Angeles. Los Angeles, CA: Hamilton, Rabinovitz and Alschuler, Inc.

Hopper, K. (1990). Public Shelter as 'a Hybrid Institution': Homeless Men in Historical Perspective. Journal of Social Issues, 46, 13–19.

Hopper, K. (1991). Homelessness Old and New: The Matter of Definition. Housing Policy Debate, 2, 757–814.

Iachan, R. (1989). Sampling in Time and Space. Proceedings of the Section on Survey Research Methods of the American Statistical Association, 636–640.

Katzman, R., Brown, T., Fuld, P., Peck, A., Schecter, R., and Schimmel, H. (1983). Validation of Short Orientation-Memory-Concentration Test of Cognitive Impairment. American Journal of Psychiatry, 140, 734–739.

National Institute on Drug Abuse (NIDA) (1991). Homeless and Transient Population Study Training Manual. Technical document under NIDA contract #271-89-8340, Washington, DC, Metropolitan Area Drug Study. Research Triangle Park, NC: Research Triangle Institute.

National Institute on Drug Abuse (NIDA) (1993). Prevalence and Treatment of Drug Use and Correlated Problems in the Homeless and Transient Population: 1991. Dennis, M.L., Iachan, R., Thornberry, J.P., Bray, R.M., Packer, L.E., and Bieler, G.S. Final report under NIDA contract No. 271-89-8340, Washington, DC, Metropolitan Area Drug Study. Rockville, MD: NIDA.

Ringwalt, C., and Iachan, R. (1990). Design Summary of Proposed Study to Estimate the Characteristics of Runaway and Homeless Youths. Research Triangle Park, NC: Research Triangle Institute.

Rossi, P.H. (1989). Down and Out in America: The Origins of Homelessness. Chicago, IL: The University of Chicago Press.

Rossi, P.H., Fisher, G.A., and Willis, G. (1986). The Condition of the Homeless of Chicago. Amherst, MA, and Chicago, IL: Social and Demographic Research Institute and NORC.

Santiago, J.M., Bachrach, L.L., Berren, M.R., and Hannah, M.T. (1988). Defining the Homeless Mentally Ill: A Methodological Note. Hospital and Community Psychiatry, 39, 100–101.

Vernez, G., Burnam, M.A., McGlynn, T.S., and Mittman, B.S. (1988). Review of California's Program for the Homeless Mentally Disabled. Final report under contract no. 86-77219 to the California Department of Mental Health. Santa Monica, CA: Rand Corporation (R-3631-CDMH).