3. SAMPLE DESIGN, SELECTION, AND MANAGEMENT

The National Survey of Veterans (NSV 2001) was intended to provide estimates for the entire non-institutionalized U.S. population of veterans, as well as for veteran population subgroups of special interest to the Department of Veterans Affairs (VA). The subgroups of primary interest were the seven health care enrollment priority groups. The VA was also particularly interested in data for female, African American, and Hispanic veterans. In addition, the survey was required to provide information needed for major initiatives that would have a direct effect on veterans, such as benefit eligibility reform and health care benefit reform. The sample design had to accommodate these policy issues.

3.1   Sample Design

The VA desired to obtain 95 percent confidence intervals of ±5 percent or smaller for estimates of proportion of 0.5 for each of the veteran population subgroups. The resulting design called for 20,000 interviews to be completed by random selection of veterans. As discussed later in this section, we evaluated a number of alternative sample design options and adopted a dual frame design consisting of a random digit dialing sample (RDD Sample) and a List Sample. The cost-variance optimization resulted in sample allocation of 13,000 completed interviews with random digit dialing method and 7,000 completed interviews from the List Sample. The List Sample design used the VHA Healthcare enrollment file and the VBA Compensation and Pension (C&P) file to construct the sampling frame. The VA administrative files alone could not be used for the sample design because the coverage from these files was only about 21 percent.

Veterans living in institutions were included in the survey target population only if they were in the institution for less than 6 months and also had a principal residence elsewhere. Such veterans were included in the survey as part of the RDD Sample only. Although the list frame contained institutionalized veterans, they were not interviewed as part of the List Sample because these would have to be screened for eligibility. Veterans living abroad and in the territories were also excluded from the survey target population. Therefore, a veteran sampled from the list frame was not eligible for the survey if the address was outside of the continental United States and Puerto Rico.

Allocation of Sample across Priority Groups

According to year 2000 projections of the veteran population provided by the VA, approximately 25 million veterans were living across the country. The VA manages its provision of health care services by assigning veterans who enroll in their health care system to one of seven health care enrollment priority groups, outlined as follows:

- <u>Priority 1</u>. Veterans with service-connected[1] conditions rated 50 percent or more disabling.

- <u>Priority 2</u>. Veterans with service-connected conditions rated 30 to 40 percent disabling.

- <u>Priority 3</u>. Veterans who are former POWs. Veterans with service-connected conditions rated 10 to 20 percent disabling. Veterans discharged from active duty for a condition that was incurred or aggravated in the line of duty. Veterans awarded special eligibility classification under 38 U.S.C., Section 1151.

- <u>Priority 4</u>. Veterans who receive increased pension based on a use of regular aid and attendance or by reason of being permanently housebound and other veterans who are catastrophically disabled.

- <u>Priority 5</u>. Veterans with nonservice-connected and veterans with noncompensated service-connected conditions who are rated zero percent disabled, and whose income and net worth are below an established threshold.

- <u>Priority 6</u>. All other eligible veterans who are not required to make co-payments for their care. This includes:

- World War I and Mexican Border War veterans;

- Veterans solely seeking care for disorders associated with exposure to a toxic substance, radiation, or for disorders associated with service in the Persian Gulf; and

- Veterans with service-connected conditions who are rated zero percent disabled but who are receiving compensation from the VA.

- <u>Priority 7</u>. Veterans with nonservice-connected disabilities and veterans with noncompensated service-connected conditions who are rated zero percent disabled, and who have income or net worth above the statutory threshold and who agree to pay specified co-payments.

---

[1] The term "service-connected" refers to a VA decision that the veteran's illness or injury was incurred in, or aggravated by, military service.

The distribution of the total veteran population across the seven priority groups is given in Table 3-1. Further, the law defines two eligibility categories: mandatory and discretionary. Priority groups 1 through 6 are termed as mandatory, whereas priority group 7 is termed as discretionary.

Table 3-1.   Distribution of total veteran population across priority groups

	Mandatory						Discretionary
Priority group	1	2	3	4	5	6	7
Percent of total	2.31	2.06	5.01	0.73	29.96	0.34	59.59

Note: These distributions do not reflect actual veteran health care enrollments. These distributions were provided by VA analysts as estimates of what the veteran population would look like if it was segmented into the seven priority groups.

Three Approaches to Sample Allocation

The VA required that the sample design produce estimates of proportions for veterans belonging to each of the seven priority groups and for female, Hispanic, and African American veterans. Therefore, different sampling rates had to be applied to the seven healthcare enrollment priority groups. In particular, priority groups 4 and 6 had to be sampled at relatively higher sampling rates to produce estimates with the required levels of reliability.

We considered three approaches to allocate the total sample across the seven priority groups: (1) equal allocation, (2) proportional allocation, and (3) compromise allocation.

Approach I – Equal Allocation

Under this approach, the sample is allocated equally to each of the seven priority groups. The equal allocation approach achieves roughly the same reliability for the priority group estimates of proportions. In other words, it achieves almost the same coefficient of variation for all priority group estimates. Because the veteran population varies across priority groups, choosing this approach would have meant that the selection probabilities of veterans would have also varied across priority groups. As a result, the variation between the sampling weights would have been very large and would have resulted in large variances for the national level estimates. We therefore did not choose this allocation because it would not have been very efficient for the national level estimates.

Approach II – Proportional Allocation

For this approach, the sample is allocated to the priority groups based on the proportion of the veteran population that each priority group represents. Under the proportional allocation approach, the priority groups with larger veteran populations would have received the larger share of the sample. In particular, priority group 7 would have received a very large sample, while the sample sizes for priority groups 4 and 6 would have been too small to produce reliable survey estimates. The proportional allocation would be the most efficient allocation for the national level estimates because the probabilities of selection are the same for all veterans irrespective of the priority group. We did not choose this allocation because reliable priority group estimates would only have been possible for the three largest groups (priority groups 3, 5, and 7).

Approach III – Compromise Allocation

As the name implies, the compromise allocation is aimed at striking a balance between producing reliable priority group estimates (*Approach I*) and reliable national level estimates (*Approach II*). A number of procedures are available to achieve this compromise. The actual procedure to be applied depends on the exact survey objectives. The simplest and most commonly used allocation is the so-called "square root" allocation. Under this allocation, the sample is allocated to the priority groups proportional to the square root of the population of the priority groups. Under the "square root" allocation, the sample is reallocated from very large priority groups to the smaller priority groups as compared with what would have been under the proportional allocation. A more general compromise allocation is the "power allocation" discussed by Bankier (1988) under which the sample is allocated proportional to x^{l}, where x is the measure of size and the parameter l can take values between zero and 1. The value $l = \frac{1}{2}$ corresponds to the "square root allocation." The two extreme values of l give the "equal allocation" and the "proportional allocation." In other words, $l = 0$ corresponds to *Approach I*, which is "equal allocation" and $l = 1$ corresponds to *Approach II*, which is "proportional allocation." Kish (1988) has also considered a number of compromise allocations including the "square root" allocation.

Because we were interested in both national level estimates and the estimates for each of the priority groups, we used the "square root" compromise allocation to allocate the sample across the seven priority groups. The sample allocation across the seven priority groups under the "square root" compromise allocation is shown in Table 3-2. The sample allocation under the proportional allocation is identical to the distribution of the veteran population across priority groups (Table 3-1) and that under the equal allocation would assign 14.3 percent of the sample to each of the priority groups. In order to achieve the "square root" allocation for minimum cost we chose a dual frame design.

Table 3-2.   Allocation of NSV 2001 sample across priority groups under "square root" allocation

Priority group	1	2	3	4	5	6	7
Percent of sample	7.66	7.25	11.29	4.32	27.61	2.92	38.95

Dual Frame Sample Design

Although it would have been theoretically feasible to select an RDD Sample with "square root" allocation of the sample across priority groups, such a sample design would have been prohibitively expensive. The RDD Sample design is an Equal Probability Selection Method (epsem) design, meaning that all households are selected with equal probability. Thus, a very large RDD Sample would have to be selected in order to yield the required number of veterans in priority group 6, the priority group with the smallest proportion of veterans. The alternative was to adopt a dual frame approach so that all of the categories with insufficient sample size in the RDD Sample could be directly augmented by sampling from the VA list frame. The corresponding survey database would be constructed by combining the List and the RDD Samples with a set of composite weights. This approach allowed us to use both samples to achieve the desired level of precision for subgroups of interest to the VA.

RDD Sample Design

We used a list-assisted RDD sampling methodology to select a sample of telephone households that we screened to identify veterans. This methodology was made possible by recent technological developments (Potter et al., 1991, and Casady and Lepkowski, 1991 and 1993). In

list-assisted sampling, the set of all telephone numbers in an operating telephone exchange is considered to be composed of 100-banks. Each 100-bank contains the 100 telephone numbers with the same first eight digits (i.e., the identical area code, telephone exchange, and first two of the last four digits of the telephone number). All 100-banks with at least one residential telephone number that is listed in a published telephone directory, known as "one-plus listed telephone banks," are identified. We restricted the sampling frame to the "one-plus listed telephone banks" only and then selected a systematic sample of telephone numbers from this frame. Thus, the RDD sampling frame consisted of all the telephone numbers in the "100-banks" containing at least one *listed* telephone number.

The nonlisted telephone numbers belonging to "zero-listed telephone banks" were not represented in the sample. However, nonlisted telephone numbers that appeared by chance in the "one-plus listed telephone banks" were included in the list-assisted RDD sampling frame.

Therefore, the list-assisted RDD sampling approach has two sources of undercoverage. The first is that nontelephone households are not represented in the survey. The second is the loss of telephone households with unlisted telephone numbers in the banks having no listed telephone numbers, known as "zero-listed telephone banks." Studies have been carried out on these potential losses, and the undercoverage from the two sources is estimated to be only about 4 to 6 percent (Brick et al., 1995). As discussed in Chapter 6, an adjustment to correct for the undercoverage was applied by use of a raking procedure with estimated population counts from the Census 2000 Supplementary Survey (C2SS) conducted by the U.S. Bureau of the Census.

List Sample Design

The VA constructed the list frame from two VA administrative files, the 2000 VHA Healthcare enrollment file and the 2000 VBA Compensation and Pension (C&P) file. The files were crossed against each other, and a single composite record was created for each veteran by matching the Social Security numbers. The list frame included information about the priority group to which each veteran belonged. Table 3-3 lists the total veteran population and the percentage of population represented by the list frame for each of the priority groups.

Table 3-3.   Percentage of veterans in the VA files by priority group

Priority group	Veteran population (thousands)	Percentage of veterans in the list frame
1	577.5	100.0
2	516.4	100.0
3	1,254.1	100.0
4	183.6	94.7
5	7,501.4	25.5
6	83.8	100.0
7	14,920.3	5.9
All veterans	25,037.1	21.6

As observed in Table 3-3, the two largest priority groups (groups 5 and 7) have very low coverage of the veteran population in the list frame, whereas four out of the remaining five priority groups (groups 1, 2, 3, and 6) have 100 percent coverage. The list frame provides almost 95 percent coverage for priority group 4 (the second smallest priority group). This feature of the list frame was advantageous for the dual frame sample design because the sample could be augmented from the list frame for the smaller priority groups. The VA lists covered 21.6 percent of the overall veteran population including the priority group 7 veterans. Because of the very large proportion of priority group 7 population, no List Sample was required to augment this group of veterans. After excluding priority group 7 veterans, the list frame contained a total of over 4.5 million veterans, accounting for 44.7 percent of the veteran population belonging to the mandatory health care groups (priority groups 1 through 6).

The list frame was stratified on the basis of priority group (groups 1 through 6) and gender. Thus, the veterans on the list frame were assigned to one of 12 design strata and a systematic sample of veterans was selected independently from each stratum.

Allocation of Sample to List and RDD Frames

Because it was less costly to complete an interview with a case from the List Sample than the RDD Sample, the goal was to determine the combination of List and RDD Sample cases that would achieve the highest precision at the lowest cost. The higher RDD unit cost was due to the additional screening required to identify telephone households with veterans.

The largest proportion of veterans is in priority group 7, which accounts for 59.6 percent of the total veteran population. The proposed "square root" sample allocation scheme meant that we would allocate 38.9 percent of the total sample to priority group 7 veterans. Let n be the total sample size and a be the proportion of the total sample that will be allocated to the RDD frame. Then the expected RDD sample in priority group 7 would be $0.596 \times a \times n$. The sample required for priority group 7 under the square root allocation was equal to $0.389 \times n$. Because no sample augmentation from the list frame was required for priority group 7 the RDD sample in priority group 7 must be equal to the sample required for the priority group, i. e. $0.596 \times a \times n = 0.389 \times n$, which gives $a = 0.653$. Thus, we needed to allocate 65.3 percent of the total sample to the RDD frame. Any smaller proportion allocated to the RDD frame would have had an adverse impact on the reliability of the estimates, and a larger RDD proportion would have increased the cost. Thus, 65.3 percent was the optimum allocation that minimized the cost while achieving square root allocation of the total sample across priority groups. The proportion was rounded to 65 percent for allocation purposes, that is, 65 percent of the total sample was allocated to the RDD frame.

The NSV 2001 cost assumptions were based on the previous RDD studies and the assumption that about one in four households would be a veteran household. We determined from these assumptions that it would be 1.3 times as expensive to complete an interview from an RDD household veteran as compared with a List Sample veteran. As discussed later in this chapter, a number of alternate sample designs were evaluated for the total cost and the design effects for various veteran population subgroups of interest.

Sample Size Determination

The decision on the sample size of completed extended interviews was guided by the precision requirements for the estimates at the health care priority group level and for the population subgroups of particular interest (namely, female, African American, and Hispanic veterans). The 95 percent confidence interval for a proportion equal to 0.5 was required with 5 percent or smaller confidence interval half-width for these population subgroups.

The sample size required for the 95 percent confidence interval with desired half-width (w) for a proportion of $p=0.5$ can be determined by solving the following equation for the sample size n_c

$$1.96\sqrt{\left(\frac{0.25}{n_c}\right)\times(deff)} = w,$$

where *deff* is the design effect for the corresponding survey estimate. As discussed later in this chapter, the *deff* of a complex sample design is the ratio of the variances under the complex design and the simple random sample design with the same sample sizes. For example, the sample size for each priority group would be 768 for 95 percent confidence interval with 5 percent margin of error for a sample design with *deff* equal to 2.0. In order to assign a sample of 768 completed interviews to priority group 6 (the priority group with smallest proportion of veterans), while maintaining "square root" allocation across priority groups, we would have to complete more than 26,000 interviews. This sample size was larger than VA was prepared to select and it was decided to reallocate the sample across priority groups by departing slightly from the proposed "square root" allocation and accepting larger sampling errors for some veteran population subgroups. As a result, the sample size of 20,000 completed interviews was sufficient to satisfy the new precision requirements.

Alternative Sample Design Options

We evaluated six sample design options with respect to cost and design efficiency for a fixed total sample of 20,000 completed interviews. Two of the sample designs were based on RDD sampling alone, whereas the remaining four designs were based on a dual frame methodology using RDD and list sampling. For each of the sample designs considered, we compared the coefficients of variation of the national estimates and those of a number of veteran population subgroups as well as the corresponding design effects. The design effects were computed to evaluate the efficiency of each of the alternative sample designs. We also obtained the cost estimates for the alternative sample designs using linear cost models incorporating screener and extended interview unit costs.

Sample Design A

This sample design is based on RDD sampling of the number of veteran households that would yield a sample of 20,000 completed extended interviews. The sample sizes across the seven veteran health care priority groups are random, and the expected sample sizes would be distributed according to their proportion in the population. Similarly, for the population subgroups of particular interest (female, African American, and Hispanic veterans), the sample sizes are also random and the expected sample sizes would be distributed in proportion to the population sizes of the respective subgroups.

Sample Design B

This is also an RDD Sample design but the sample is allocated to the priority groups according to the "square root" allocation scheme. The fixed sample sizes across priority groups are achieved through screening. The number of veteran households to be screened is determined by the sample size allocated to priority group 6, the smallest priority group, with 0.34 percent of the veteran population. The resulting sample design is a stratified sample design with "square root" allocation of the sample across the seven priority groups.

Sample Design C

This is the dual frame sample design discussed earlier, with an RDD Sample of 13,000 and a List Sample of 7,000 completed extended interviews. The list frame sample design is a stratified sample design. The first level of stratification is on the basis of priority groups 1 through 6. As noted previously, no List Sample is allocated to priority group 7. The List Sample is allocated across the remaining six priority groups to achieve the "square root" allocation of the total sample (RDD and List) across the seven priority groups. The next level of stratification is by gender within each priority group and the sample is allocated so that the sampling rate for female veterans is twice that for male veterans. The stratification by gender allowed us to increase the sample size for female veterans by sampling at a higher rate. This strategy could not be adopted for Hispanic and African American veterans because the variables to identify race/ethnicity were not available on the VA files used to construct the list frame.

Sample Design D

This sample design is essentially the same as sample design C but the List Sample is reallocated across the six priority groups by oversampling priority groups 6 and 4 and correspondingly reducing the sample size for priority group 5. Priority group 7 veterans are not selected from the list frame.

Sample Design E

This is a dual frame design with an RDD Sample of 10,000 and a List Sample of 10,000 completed extended interviews. To achieve the "square root" allocation of the total sample across the priority groups, the List Sample must be allocated to priority group 7 veterans as well. As before, the List Sample is a stratified sample and the sampling rate for female veterans is twice the sampling rate for male veterans within each priority group.

Sample Design F

This is also a dual frame design with an RDD Sample of 15,000 and a List Sample of 5,000 completed extended interviews. To achieve the "square root" allocation of the total sample across the priority groups, the RDD Sample must be screened for priority groups. The List Sample is allocated to priority groups 1 through 6. As in the case of other dual frame designs, the List Sample design is a stratified sample design, and the sampling rate for female veterans is twice the sampling rate for male veterans within each priority group.

Efficiencies of the Alternative Sample Designs

To evaluate the precision of the survey estimates, we computed the standard errors of these estimates, where the standard error of an estimate is defined as the square root of the variance of the estimate. The ratio of the estimated standard error to the survey estimate itself, called the coefficient of variation (cv) of the survey estimate can also be used to evaluate a sample design.

Another way to evaluate the efficiency of a sample design and the procedure used to develop the survey estimates is by using the design effect. Design effect is defined as the ratio of the variance of an estimate for a complex sample design and the variance of the estimate under the simple random sample design with the same sample size. Kish (1965) introduced the concept of design effect to deal with complex sample designs involving stratification and clustering. Stratification generally leads to a gain in efficiency over simple random sampling, but clustering usually leads to deterioration in the efficiency of the sample design due to positive intracluster correlation among units in the cluster. To determine the total effect of any complex design on the sampling variance in comparison to the alternative simple random sample design, the design effect *(deff)* is defined as

$$Deff = \frac{sampling \quad variance \quad of \quad a \quad complex \quad sample \quad design}{sampling \quad variance \quad of \quad simple \quad random \quad sample \quad design}.$$

We used the design effects for various survey estimates, including the estimates for priority groups, to evaluate the alternative sample designs that we considered for the NSV 2001. The smaller the design effect the more efficient is the sample design without taking the cost into consideration.

We also computed the cv of the survey estimates to check the precision requirements for the survey estimates. The precision requirement was specified in terms of margin of error of the 95 percent confidence interval, which is given by 1.96 times the cv of the estimate. Table 3-4 provides the design effects and the cv of estimates of proportions equal to 0.5 for various population subgroups, including the priority groups for the alternative sample design options.

The following sections of this chapter discuss the comparative costs of the six alternative sample design options and the cost-variance efficiencies of the alternative sample designs.

Cost Comparisons

The choice of a survey design involves comparing costs and the design efficiencies in terms of sampling variances. To estimate the total survey costs for the six alternative sample

designs, we used a linear cost model. A general disposition of the total sampling cost for each of the alternative sample designs can be described as follows.

Table 3-4.   Design effects and coefficients of variation (cv) for various veteran population subgroups for the alternative sample designs

| Characteristic | Design Effects | | | | | |
	Design A	Design B	Design C	Design D	Design E	Design F
All veterans	1.00	1.27	1.48	1.48	1.92	1.30
Priority 1	1.98	1.00	1.13	1.13	1.10	1.21
Priority 2	1.98	1.00	1.12	1.12	1.10	1.20
Priority 3	1.95	1.00	1.18	1.18	1.14	1.32
Priority 4	1.99	1.00	2.09	2.47	2.42	1.97
Priority 5	1.70	1.00	2.07	1.92	2.62	2.02
Priority 6	2.00	1.00	1.04	1.04	1.04	1.07
Priority 7	1.40	1.00	1.39	1.39	1.75	1.21
Male	1.05	1.34	1.52	1.52	1.95	1.35
Female	1.95	2.45	2.98	2.96	4.15	2.57
African American	1.92	2.44	2.50	2.52	3.20	2.36
Hispanic	1.96	2.50	2.55	2.57	3.26	2.41

| Characteristic | Coefficients of Variation | | | | | |
	Design A	Design B	Design C	Design D	Design E	Design F
All veterans	0.35	0.40	0.43	0.43	0.49	0.40
Priority group 1	3.30	1.28	1.36	1.36	1.34	1.40
Priority group 2	3.44	1.31	1.38	1.38	1.37	1.43
Priority group 3	2.19	1.05	1.14	1.14	1.12	1.20
Priority group 4	6.84	1.83	2.65	2.48	2.85	2.57
Priority group 5	0.86	0.68	0.97	0.98	1.10	0.96
Priority group 6	9.44	2.15	2.16	1.80	2.19	2.23
Priority group 7	0.54	0.56	0.66	0.66	0.74	0.62
Male	0.37	0.42	0.45	0.45	0.51	0.42
Female	2.18	2.47	2.43	2.44	2.77	2.32
African American	1.71	1.93	1.95	1.96	2.21	1.89
Hispanic	2.47	2.80	2.82	2.84	3.19	2.74

Selection of an RDD Sample considered for the NSV 2001 involved two interviewing processes: a household screening interview and an extended interview. Depending on an allocation method adopted in the sample design, the screening interview would be administered to screen only for the veteran households (level I screening) or for the priority groups (level II screening). Under the proportional allocation, sample sizes to be realized across seven priority groups will be random variables, but their expected sample sizes will be

proportional to the population sizes of the priority groups. The interviewer would need to screen only for the veteran households (level I screening) to get the required sample size, say n_{RDD}, for the RDD Sample. The RDD Sample for sample design A, C, D, and E were classified into this category, and the general expression for such screening costs is given by

$$C_{Screen\_I} = n_{RDD} \times C_1,$$

where C_1 is the unit cost for the level I screening.

On the other hand, the square root allocation used in the RDD Sample for sample design B necessitated an additional step to screen for priority group (level II screening) as well. The RDD Sample for sample design F also must go through this step because the corresponding RDD Sample must be screened for the priority groups. That is, once the interviewer found a veteran household, he or she would need to further screen for the priority group. Determination of priority group that a veteran belongs to would require asking questions on income, assets, debts, number of dependents, and disability rating. Consequently, the corresponding unit cost C_2 for the level II screening was much larger than the unit cost C_1 for the level I screening. Also, to obtain the designated sample sizes, the screening process would need to continue until the designated sample size for priority group 6 (the smallest category) is obtained. Thus, the number of telephone numbers to be screened, denoted by m_{RDD}, would be much larger than the corresponding total number of telephone numbers to be sampled for the design, n_{RDD}. The higher unit screening cost C_2 for level II screening and the larger number m_{RDD}, of households to be screened would result in a very large total screening cost. The total screening cost for level II screening is given by

$$C_{Screen\_II} = m_{RDD} \times C_2.$$

On completing the screening interview, the extended interview would be administered to the fixed number, n_{RDD}, of veterans. Letting $C_0^{(RDD)}$ denote the unit cost of the RDD extended interview, the total cost for the RDD extended interview is

$$C_{RDD-Extended} = n_{RDD} \times C_0^{(RDD)}.$$

This equation applies to all RDD Samples.

For the List Sample, relevant information for selecting veterans is available at the outset from the frame itself. Thus, no screening cost is incurred but there is a cost associated with tracking and tracing the List Sample veterans. Moreover, the average time to administer the extended interview to List Sample veterans would be higher that that for the RDD Sample veterans. Therefore, the unit cost of the List extended interview would be higher than the RDD extended interview unit cost. The extended interview would need to be administered to the List Sample cases for all of the alternative sample designs using a dual frame methodology. If we denote by $C_0^{(List)}$ the unit cost of the List Sample extended interview then the total interview cost of the List extended interview is

$$C_{List-Extended} = n_{List} \times C_0^{(List)}.$$

Under a linear cost model, total sampling cost based on the above disposition can be obtained as the sum of the relevant components. For example, $C_{Total} = C_{Screen\_I} + C_{RDD-Extended}$ is the total survey cost for the sample design A, $C_{Total} = C_{Screen\_II} + C_{RDD-Extended}$ is the total survey cost for sample design B, $C_{Total} = \left(C_{Screen\_I} + C_{RDD-Extended} \right) + C_{List-Extended}$ is the total survey cost for sample design C, and so on. Table 3-5 provides the relative total sampling costs for the alternative sample designs, where we have set the total cost for sample design A equal to 100 to allow a standard comparison across sample designs.

Table 3-5.   Cost comparison of the alternative sample designs

Sample design option	Cost relative to option A
A	100
B	519
C	93
D	93
E	89
F	104

As expected, the cost of the sample designs C and D are the same because of the same sample allocation between the list and RDD frames. On the other hand, sample design B is very expensive due to higher screening cost (level II screening).

Cost-Variance Efficiency of the Alternative Sample Designs

We considered six sample designs to meet the survey objective, which was to obtain sufficient data about a cross section of veterans for each health care priority group, while at the same time obtaining reliable survey estimates at the national level. Reliable survey estimates were also required for female veterans and for Hispanic and African American veterans. The design effects and coefficients of variation for the six sample design options were obtained and these are given in Table 3-4 for various subgroups of the veteran population. The reliabilities of the survey estimates of proportions equal to 0.5 for various domains of interest were considered for evaluating the alternative sample design options.

Although the design effects would always be less than 2 for sample design A, this design does not satisfy the survey objectives due to very small sample sizes for smaller priority groups. The design effect for sample design A is equal to $(2 - d)$ when estimating a proportion equal to 0.5 for a population domain of size d relative to the entire population. On the other hand, sample design B would satisfy the survey objectives but such a design would not be feasible because of very high screening costs. Hence, neither of the two RDD Sample designs is suitable for the NSV 2001, and we used a dual frame sample design with the allocation of the total sample to the RDD and list frames as discussed earlier in this chapter.

The cost of sample design C reduces to 93 percent of sample design A because it is a dual frame design. The precision requirement is also satisfied for most veteran population subgroups. It turned out that for sample design C, the highest design effects were for priority groups 4 and 5 because coverage of these two groups by the list frame was less than 100 percent. In spite of the high design effect for priority group 5, the precision requirement was satisfied because of the larger sample size. The sample size required to satisfy the desired reliability for the priority group 4 estimates could not be achieved by the proposed "square root" sample allocation across priority groups with the total sample size of 20,000 completed interviews. Although the proportion of veterans in priority group 6 was less than that in priority group 4, the precision requirement for priority group 6 was satisfied because of its 100 percent overlap with the list frame.

The design effects for female, African American, and Hispanic veterans were also larger than 2. The female veterans account for 5.4 percent of the total veteran population, and Hispanic and African Americans are respectively 4.0 percent and 8.2 percent of the total veteran population. The precision requirements for the estimates for female veterans were achieved

through oversampling the list frame. The higher sampling rate for female veterans increased the design effects slightly for the national level estimates because of increased variability in the survey weights. The precision requirements for the estimates for Hispanic veterans could not be satisfied. The List Sample size for Hispanic veterans could not be increased because the variables to define race/ethnicity were not available on the list frame. The precision requirement for the estimate for African Americans was met because of a larger sample size.

We reallocated the sample as proposed under sample design D to improve the precision of the two smallest priority groups (groups 6 and 4). Under sample design D we reallocated the sample from priority group 5 to priority groups 4 and 6 from the allocation under sample design C. Although the precision requirement for priority group 6 was satisfied because the list frame covered this group completely, the List Sample was also reallocated from priority group 5 to priority group 6. Under sample design D the 95 percent confidence interval half-width would be wider than 5.0 percent for Hispanic veterans only. Although the design effect for priority group 4 increased, the precision level improved because of an increase in the sample size due to sample reallocation. The increase in the design effect for priority group 4 resulted from fixed RDD variance for the nonoverlap part. Priority group 6 has 100 percent coverage by the list frame, and hence the design effect does not change if the List Sample size increases. Overall, when the List Sample is reallocated, the precision of the estimates for both priority groups 4 and 6 improves without significantly deteriorating the precision of the estimate for priority group 5. The precision of the estimate for Hispanic veterans decreases somewhat, but the effect is almost negligible (the design effect for Hispanic veterans increased from 2.55 under sample design C to 2.57 under sample design D). The survey cost for sample design D is the same as under sample design C due to the same sample allocation between the list and the RDD frames.

The other two dual frame designs (sample designs E and F) differ from sample design C in sample allocation between the RDD and list sampling frames. Under design E, the cost is reduced only slightly (89 for design E versus 93 for design C) by over-allocating the sample to the list frame, but the reliability of the survey estimates is affected for several categories of estimates. Sample design F achieves better precision levels for some of the categories and worse for others, but the overall cost increases because of higher screening costs (level II screening). Sample design F requires screening for priority group in order to achieve "square root" allocation of the sample at the priority groups. The cost of sample design F is 104 as compared with 93 for sample designs C and D. Therefore, sample design D provides a solution that satisfies the survey objectives of producing reliable estimates and controlling the overall cost of the survey.

The sampling parameters of sample design D (sample allocation and sample sizes) are given in Table 3-6. The table also gives the effective sample size, defined as the total sample size divided by the design effect. The minimum effective sample size must be 384 in order to achieve the required 5 percent half-width for 95 percent confidence interval of the estimate of proportion equal to 0.5. Thus, for sample design D, the only veteran population subgroup for which the precision requirement could not be met was Hispanics.

Table 3-6.   Sample allocation for sample design D

Characteristic	RDD	Sample size List	Total	Design Effect	Effective sample size
All veterans	13,000	7,000	20,000	1.48	13,489
Priority group 1	295	1,240	1,535	1.13	1,357
Priority group 2	271	1,199	1,470	1.12	1,308
Priority group 3	661	1,636	2,296	1.18	1,939
Priority group 4	69	931	1,000	2.47	405
Priority group 5	3,731	1,231	4,962	1.92	2,589
Priority group 6	36	764	800	1.04	773
Priority group 7	7,937	0	7,937	1.39	5,712
Male	12,338	6,419	18,757	1.52	12,344
Female	662	581	1,243	2.96	420
African American	1,066	574	1,640	2.52	650
Hispanic	520	280	800	2.57	311

3.2   Sample Selection

The samples from the list and RDD frames were selected independently. The RDD Sample consists of a sample of telephone households, and the List Sample consists of veterans sampled from the VA list frame. This section describes sampling procedures for each of the two components.

List Sample Selection

The List Sample is a stratified sample with systematic sampling of veterans from within strata. The strata were defined on the basis of priority group and gender. The first level of stratification was by priority group and then each priority group was further stratified by gender. Thus, the sample had 12 strata (priority group by gender).

Under the assumption of an 80 percent response rate to the main extended interview, a List Sample of about 8,750 veterans was anticipated to yield 7,000 complete interviews. We also decided to select an additional 50 percent reserve List Sample to be used in the event that response rates turned out to be lower than expected. Therefore, the total sample size selected from the list frame was 13,125 veterans. With the systematic sampling methodology, we achieved a total sample of 13,129 veterans from the list frame, out of which a sample of 4,377 veterans was kept as a reserve sample.

The allocation of the List Sample to the six priority groups, in combination with the RDD Sample, corresponded to sample design D. Female veterans were sampled at twice the rate as male veterans while keeping the List Sample size fixed at 13,129. Table 3-7 provides the allocation of the List Sample to the 12 sampling strata.

Table 3-7.   Allocation of List Sample to sampling strata

Priority group	Gender	Stratum	Sample Size
1	Male	11	2,082
	Female	12	242
2	Male	21	2,027
	Female	22	224
3	Male	31	2,798
	Female	32	270
4	Male	41	1,637
	Female	42	110
5	Male	51	2,127
	Female	52	182
6	Male	61	1367
	Female	62	63
Total sample			13,129

Based on the sizes of the 12 sampling strata on the list frame and the allocated sample size for each stratum given in Table 3-7, we used a systematic random sampling procedure within each stratum to select List Sample veterans.

We also determined release groups by assigning sequential numbers to each veteran in the List Sample starting with 1. The release groups were defined for sample management purposes as discussed later in this chapter. The List Sample was divided into 15 release groups by assigning veterans numbered i, $i+15$, $i+30$, $i+45$, etc. to the ith release group, where $i =1,2,...,15$. The first 4 release groups contained 876 veterans each and the remaining 11 release

groups contained 875 veterans each. From those 15 release groups, we selected two sets (waves) consisting of 5 release groups each for the March 2001 and the May 2001 List Sample releases. These were selected by using two sequential systematic samples of release groups from the 15 release groups. We used the remaining 5 release groups as a reserve sample to be released only if the actual response rates turned out to be lower than the assumed response rate of 80 percent.

Although the cooperation rates were quite high, the List Sample targets could not be met with the initial two waves of the List Sample. This happened because of out-of-scope sample cases (e.g., deceased, institutionalized veterans) and sampled cases that could not be traced. Thus, we released the final wave, which also consisted of 5 release groups, in June 2001.

RDD Sample Selection

National RDD Sample

We selected the RDD Sample of households using the list-assisted RDD sampling method. This method significantly reduces the cost and time involved in such surveys in comparison to dialing numbers completely at random. The general approach we employed was a two-stage sampling procedure in which we initially selected a sample of telephone numbers and successfully screened for households with veterans.

Using the list-assisted sampling methodology, we selected a random sample of telephone numbers from "one-plus listed telephone banks." This list-assisted RDD sampling methodology is implemented in the GENESYS sampling system, which employs a single-stage equal probability sampling method to select the telephone numbers. The "one-plus listed telephone banks" are initially sorted by geographic variables, such as state, metropolitan, and nonmetropolitan, and also by area codes and five digit prefixes. These sorts construct the sampling frame. The frame is then divided into implicit strata (almost) equal in size while preserving the sort ordering. The total number of such implicit strata is the same as the desired sample size. Then a single telephone number is selected independently from within each implicit stratum.

Based on propensity estimates from the 1992 NSV RDD Sample, we estimated that we needed a sample of 135,440 telephone numbers to obtain 13,000 completed extended interviews for the RDD component of the sample. Our assumptions were:

- Residential numbers – 60 percent;

- Response to screening interview – 80 percent;

- Households with veterans – 25 percent; and

- Response to extended interview – 80 percent.

The sample yields at various steps during the RDD Sample selection are given in Figure 3-1.

Telephone Numbers	181,000
Residency Rate	45%
Identified Residential Households	81,450
Screener Response Rate	80%
Screened Households	65,160
Households with Veterans	25%
Identified Veteran Households	16,290
Extended Interview Response Rate	80%
Completed Extended Interviews	13,032

Figure 3-1. Expected RDD sample yield

From the above calculation, we determined that we required an RDD Sample of 181,000 telephone numbers to yield 13,032 completed extended interviews. We also decided to select an additional 30 percent reserve RDD Sample to be used in the event that the yield assumptions did not hold (in other words, if the response rate turned out to be lower than expected). Thus, a total of 235,300 telephone numbers were to be selected from the RDD frame but we increased the sample size to 240,000 telephone numbers so that 186,000 telephone numbers could serve as the main RDD Sample and the remaining 54,000 as the reserve sample. We selected the sample of 240,000 telephone numbers from the GENESYS RDD sampling frame as of December 2000.

To maintain better control over the RDD sampling yields, the 240,000 telephone numbers selected were divided into 80 release groups, with each group containing 3,000

telephone numbers. To determine release groups, sequential numbers starting from 1 were assigned to the telephone numbers in the RDD Sample. The sampled telephone numbers were kept in the same order in which they were originally listed in the GENESYS sampling frame to preserve the implicit stratification of the sample. The telephone number with the sequence numbers *i, i+80, i+160, i+240*, etc. constituted release group i, where $i = 1,2,...,80$. From the 80 release groups, systematic samples of 10 release groups (or waves) were selected almost every month for the sampling process. The first release (wave) contained 20 release groups because it included the pretest effort. The sample released in August contained 19 release groups or 57,000 sampled telephone numbers. Early in data collection it became clear that the sample yield assumptions were very optimistic and even the entire RDD Sample of 240,000 telephone numbers would not be sufficient to produce the required 13,000 completed extended interviews. Therefore, we decided to select a supplementary RDD Sample of 60,000 telephone numbers.

Supplementary RDD Sample

Based on the result of the interim RDD Sample yields, we selected a supplementary sample of 60,000 telephone numbers from the GENESYS RDD sampling frame as of June 2001. The supplementary sample was divided into 20 release groups, and the release groups were assigned the sequential numbers from 81 to 100. The supplementary sample of 60,000 telephone numbers and the one last release group from the initial RDD Sample were also released in August 2001.

Puerto Rico RDD Sample

No listed household information was available for Puerto Rico. As a result, we used a naïve RDD sampling approach called "RDD element sampling" (Lepkowski, 1988) instead of the list-assisted RDD method that we used for the national RDD Sample. With this methodology, all possible 10-digit telephone numbers were generated by appending four-digit suffixes (from 0000 to 9999) to known 6-digit exchanges consisting of 3-digit area code and 3-digit prefix combinations. A commercially available[2] database contained 325 Puerto Rico residential exchanges. Thus, the Puerto Rico RDD frame was constructed to have 3,250,000 telephone

---

[2] The 643 exchange numbers were obtained from the "Terminating Point Master Vertical & Horizontal Coordinates Data (abbreviated as TPM-VHCD)" file provided by Telcordia Technologies. As of December 7, 2000, 325 of them were associated with a regular (Plain Old) Telephone Service (abbreviated as POTS). No such numbers shared with paging for Puerto Rico.

numbers and a systematic sample of 5,500 telephone numbers was drawn to achieve 176 completed extended interviews.

Before sampling, the frame file was sorted by 6-digit exchange and place name (or service name).[3] This implicit stratification permitted a better representation of the population of households. Also, the required 176 completed extended interviews was determined proportionally to the ratio of the Puerto Rico population size (3,803,610) to the U.S. population size (281,421,906) as of April 1, 2000 from the size of the main RDD completed interviews (13,000). To achieve this target sample size of 176 completed extended interviews, the total of 5,500 telephone numbers was calculated by assuming the residency rate, the screener response rate, the veteran household rate, and the extended interview response rate as 20 percent,[4] 80 percent, 25 percent and 80 percent, respectively. Unlike the latter three rates, the assumed residential rate is much lower than that of the national RDD Sample. This is because RDD element sampling, unlike the list-assisted RDD sampling, does not have any information on the "one-plus listed telephone banks."

3.3 Sample Management

Conduct of the NSV 2001 required not only an effective sample design but also careful management of the entire sampling process, from creating the sampling frame to the end of data collection. Before each sampling step, project staff identified the goals, designed the process, and prepared detailed specifications for carrying out the procedures. At each stage, we carried out quality control procedures that involved checks of counts, cross-tabulations, valid values, and other measures of correct file processing, transfer, and handling. The remainder of this section describes the principal areas of sample management.

---

[3] This is a character field in the TPM-VHCD file, which identifies the general location or service of each area code and is used by many customer-billing processes to appear on bills. Refer to TPM-VHCD – Data Set/File Specification, Appendix A-1, Page 5.

[4] According to Lepkowski (1988, p. 83), fewer than 25% of all potential telephone numbers (generated by appending 4-digit random numbers to known area-prefix combinations) are assigned to a household.

Sampling Frames

List Frame

Based on the sample design and the data available from the VA source files, we developed specifications for constructing the list sampling frame. It included rules for:

- Defining each sample stratum;

- Determining the source file and variable for each data element needed to construct and stratify the frame; and

- Defining other needed data items and their sources, such as identification and contact information about the veteran.

The VA staff matched the two source files – the VHA Healthcare enrollment file and the VBA Compensation and Pension (C&P) file.

RDD Frame

The national RDD Sample was selected at two different times, and hence two different frames were used. The RDD frames, differing in the reference period as of December 2000 and June 2001, consisted of all "one-plus listed telephone banks" covering the fifty states and the District of Columbia. The GENESYS Sampling System provided the RDD sampling frames. As of December 2000, GENESYS had supplied 2,487,468 "one-plus listed telephone banks" in the frame. Thus, the size of the national RDD sampling frame was 248,746,800 (2,487,468×100). Westat obtains an updated sampling frame quarterly from GENESYS for conducting RDD surveys. The updated frame as of June 2001 contained 2,538,453 "one-plus listed telephone banks." As described earlier, the frames were sorted in geo-metro order before sampling so that the resulting RDD Samples represented the entire population of households.

Sample Release Groups

To ensure that the sample remained unbiased during the data collection process, we partitioned both the RDD and List Samples into a number of release groups so that each release group was a random sample. The sample was released to data collection staff in waves. Each of these sample waves comprised a number of release groups, which were selected at random. The

small size and independence of sample release groups gave precise control over the sample. During data collection, we monitored sample yield and progress toward our targets. When we noticed that a sufficient number of sample cases from the previous waves had been assigned final result codes, we released new waves of the sample. Tables 3-8 and 3-9 show the release dates of sample waves and the corresponding number of cases in each wave for the RDD and List Samples.

We carefully monitored the sample to ensure that we met the data collection goal of 20,000 completed extended interviews, and that the sample was representative of the veteran population across the sample waves (see Table 3-10). Tables 3-11 and 3-12 compare the sample yield distributions with respect to priority groups and the other demographic variables of age, gender, and race/ethnicity across sample waves. Similarly, Table 3-13 compares the sample yield distribution with respect to levels of education across waves and Table 3-14 compares that with respect to census regions.

Table 3-8.   RDD Sample waves

Sample wave	Date released	Number of cases
1	February 12, 2001	60,000
2	March 16, 2001	30,000
3	May 18, 2001	30,000
4	June 11, 2001	30,000
5	July 11, 2001	30,000
6	August 21, 2001	57,000
7	August 30, 2001	63,000
Puerto Rico	April 16, 2001	5,500

Table 3-9.   List Sample waves

Sample wave	Date released	Number of cases
1	April 9, 2001	4,376
2	May 14, 2001	4,377
3	July 27, 2001	4,376

In Table 3-10 the sample yield is defined as the ratio of the number of completed extended interviews and the number of sampled cases expressed as a percent. The sample yield was quite uniform over the List Sample waves but it decreased monotonically over the RDD Sample waves. The reason for decreasing sample yield for the RDD Sample was that interviewers had less calling time for the sample waves released later during the data collection. At various

times during data collection we closed earlier sample waves. Had we not done so, the difference in yields would have been even larger. Puerto Rico is not included in these analyses because the Puerto Rico sample was not released in waves.

Table 3-10. Sample yield by sample wave (RDD and List Samples)

Sample	Wave	Cases sampled (A)	Cases completed (B)	Percent yield (B/A) * 100
RDD	1	60,000	2,840	4.73
	2	30,000	1,374	4.58
	3	30,000	1,318	4.39
	4	30,000	1,307	4.36
	5	30,000	1,241	4.14
	6	57,000	2,394	4.20
	7	63,000	2,431	3.86
	Total	300,000	12,905	4.30
List	1	4,376	2,384	54.48
	2	4,377	2,317	52.94
	3	4,376	2,391	54.64
	Total	13,129	7,092	54.02

Table 3-11. Distribution of completed sample cases by health care priority group within each wave (RDD and List Samples)

| Sample | Wave | Health care priority group | | | | | | |
		1	2	3	4	5	6	7
RDD	1	2.75	3.27	7.99	0.14	16.97	11.76	57.11
	2	3.28	1.82	8.81	0.00	17.54	11.28	57.28
	3	3.19	3.11	9.18	0.00	18.06	11.46	55.01
	4	3.14	2.52	8.80	0.08	15.46	12.55	57.46
	5	3.46	1.93	9.11	0.00	18.53	10.72	56.24
	6	3.63	2.72	10.03	0.08	17.63	10.57	55.35
	7	3.54	2.26	9.13	0.08	18.92	9.91	56.15
	Total	3.27	2.60	8.98	0.07	17.63	11.09	56.36
List	1	21.48	18.62	26.13	1.47	18.54	6.63	7.13
	2	21.92	19.16	26.15	0.91	18.17	6.65	7.03
	3	22.12	18.44	27.48	1.25	18.40	5.98	6.32
	Total	21.84	18.74	26.59	1.21	18.37	6.42	6.82

The List Sample allocated to female veterans was 8.3 percent and Table 3-12 shows that the proportion of female veterans in the completed List Sample was 8.2 percent. Thus, the proportions allocated and completed were very close and the distribution by gender showed no bias. The female veterans in the List Sample would have been 9.5 percent if the proportion of female veterans on the list frame had been the same as that in the overall veteran population. In

fact, the proportion of female veterans on the list frame was about 4.3 percent as compared to 5.4 percent in the overall veteran population. Therefore, the proportion of female veterans in the completed List Sample cases was 8.2 percent because the sampling rate for female veterans was twice that for male veterans.

Table 3-12. Distribution of completed sample cases by demographic variables (age, gender, and race/ethnicity) within each wave (RDD and List Samples)

| Sample | Wave | Demographic variables | | | | | | | | |
| | | Race/Ethnicity | | | | Gender | | Age | | |
		Hispanic	Black	Other	White	Male	Female	Under 50	50-64	Over 64
RDD	1	3.73	8.63	4.08	83.56	94.96	5.04	25.92	34.65	39.44
	2	4.15	6.48	5.24	84.13	94.54	5.46	25.62	34.93	39.45
	3	4.02	7.36	4.25	84.37	94.23	5.77	24.89	33.92	41.20
	4	3.90	6.81	4.59	84.70	94.49	5.51	25.48	34.35	40.17
	5	4.83	7.57	3.79	83.80	94.20	5.80	29.90	33.76	36.34
	6	3.68	7.89	5.22	83.21	94.65	5.35	27.78	35.76	36.47
	7	3.91	7.36	4.90	83.83	94.61	5.39	26.66	34.43	38.91
	Total	3.95	7.61	4.61	83.83	94.60	5.40	26.60	34.65	38.74
List	1	4.82	12.37	5.96	76.85	91.99	8.01	24.83	34.27	40.90
	2	5.48	12.56	7.25	74.71	91.80	8.20	22.57	34.57	42.86
	3	4.27	13.01	6.27	76.45	91.59	8.41	23.00	34.96	42.03
	Total	4.85	12.65	6.49	76.02	91.79	8.21	23.48	34.60	41.92

Table 3-13. Distribution of completed sample cases by level of education within each wave (RDD and List Samples)

Sample	Wave	Education level			
		No High School	High School Diploma	Some College	Bachelor's Degree or Higher
RDD	1	13.35	31.16	28.27	27.22
	2	14.05	29.55	29.84	26.56
	3	12.67	30.05	29.51	27.77
	4	11.94	28.16	30.76	29.15
	5	14.26	30.86	28.28	26.59
	6	13.74	29.62	28.91	27.74
	7	13.57	30.52	28.92	26.98
	Total	13.41	30.14	29.06	27.39
List	1	20.64	27.60	30.03	21.73
	2	21.10	26.15	30.08	22.66
	3	19.62	27.56	31.45	21.37
	Total	20.45	27.12	30.53	21.91

Table 3-14. Distribution of completed sample cases by census region within each wave (RDD and List Samples)

Sample	Wave	Census region			
		Northeast	Midwest	South	West
RDD	1	18.70	24.08	36.90	20.32
	2	18.85	22.56	37.77	20.82
	3	21.17	22.31	37.03	19.50
	4	18.44	24.71	36.88	19.97
	5	17.32	25.38	35.86	21.43
	6	18.71	25.10	36.51	19.67
	7	16.95	25.96	35.50	21.60
	Total	18.48	24.47	36.57	20.48
List	1	15.94	20.43	43.16	20.47
	2	14.98	21.71	42.73	20.59
	3	14.39	20.83	44.42	20.37
	Total	15.10	20.98	43.44	20.47

We used chi-square statistics to test for homogeneity of distributions of the sample yield by priority group (Table 3-11), demographic variables (Table 3-12), level of education (Table 3-13), and census region (Table 3-14) across waves and found that none of the chi-square values was significant at 5 percent level of significance (Table 3-15). Thus, the time effect produced no evidence of bias across different sample waves.

Table 3-15. Chi-square values for testing homogeneity of distribution of the sample yield by various characteristics (RDD and List Samples)

Sample	Characteristic	Chi-square	Degrees of freedom	Probability
RDD	Priority	44.64	36	0.15
	Age	19.15	12	0.09
	Gender	1.53	6	0.96
	Race/Ethnicity	19.03	18	0.39
	Education	12.44	18	0.82
	Region	22.53	18	0.21
List	Priority	7.30	12	0.84
	Age	4.16	4	0.39
	Gender	0.25	2	0.88
	Race/Ethnicity	8.07	6	0.23
	Education	4.47	6	0.61
	Region	3.77	6	0.71