



**UNITED STATES DEPARTMENT OF COMMERCE**  
**Economics and Statistics Administration**  
**U.S. Census Bureau**  
Washington, DC 20233-0001

**MAY 31 2000**

MEMORANDUM FOR Chester E. Bowie  
Chief, Demographic Surveys Division

From: Alan R. Tupek *AL*  
Chief, Demographic Statistical Methods Division

Subject: Sample Design and Sampling Specifications for Blood Sample  
Survey and Buccal Swab Survey

Attached to this memorandum are specifications for selecting sample for the Blood and Buccal Swab Surveys from the pool of completed interviews in the 1999 Long-Term Care (LTC) Survey.

If you have any questions, please contact Padraic Murphy at 301-457-4294.

Attachment

cc:

L. Corder (Duke Univ.)	S. Higgins (DSD)	C. Alexander (DSMD)
B. Priboth "	D. Levins "	A. Brinkley "
	D. Nelson "	P. Murphy "
	M. Parker "	N. Penn "
	J. Pilkerton "	F. Winters "
	E. Williams "	

DSMD/VEB/PMurphy/pam/Q:\LTC\BB\_DSGN.WPD

## Table of Contents

I.	Description of the Surveys .....	1
II.	Initial Requirements and Adjustments .....	1
A.	Initial Requirements for Blood Sample Survey .....	1
B.	Initial Requirements for Buccal Swab Survey .....	1
C.	Adjustments Due to Overlap Between BSS and BUCSS Samples .....	2
D.	Adjustments to Age Group Sample Size Requirements .....	3
III.	Stratification and Selection of NSR PSUs for Blood Sample Survey .....	4
A.	Determining the Number of NSR Strata .....	4
B.	Allocating NSR Strata Between the Four Regions .....	4
C.	Final Stratification and Selection .....	4
IV.	Creating Sample File for Blood Sample Survey .....	6
A.	Selection of 80-89 Sample in BSS SR PSUs .....	6
B.	Selection of 65-79 cases in BSS SR PSUs .....	6
C.	Selection of the Sample in the BSS NSR PSUs .....	6
V.	Selection of Buccal Swab Survey Sample .....	7
A.	Sample Available for BUCSS .....	7
B.	Reallocation of BUCSS Sample Size Requirements .....	7
C.	Selecting the BUCSS Sample .....	8
VI.	Weighting Factors for Blood and Buccal Swab Surveys .....	9
VII.	Verification of Sampling for Blood and Buccal Swab Surveys .....	9
VIII.	References .....	10

## List of Tables

Table 1: Initial BSS Sample Size Requirements by Age Group .....	2
Table 2: Initial BUCSS Sample Size Requirements by Age Group .....	2
Table 3: Allocation of NSR Strata Between Regions .....	5
Table 4: BSS Sample Sizes and Expected Number of Interview .....	6
Table 5: Sample Available for the BUCSS .....	7
Table 6: Final Designated Sample Sizes and Expected Completions for the BUCSS .....	8
Table 7: Calculation of Weighting Factors for the BSS and BUCSS .....	9

## Appendices

Appendix A: Stratification and Selection of BSS NSR PSUs .....	A-1
Appendix B: List of BSS SR PSUs .....	B-1
Appendix C: Weighting Factors for BSS and BUCSS .....	C-1

## **I. Description of the Surveys**

The Blood Sample (BSS) and Buccal Swab (BUCSS) Surveys will be drawn from respondents to the 1999 Long Term Care Survey (LTC) who completed detailed interviews for that survey. There were 6,202 completed LTC detailed interviews.

The purpose of both surveys is to collect genetic information about the respondents. Each respondent selected for the BSS will be asked for a sample of their blood. Each respondent selected for the BUCSS will be asked for a sample of cells obtained by swabbing inside the mouth.

## **II. Initial Requirements and Adjustments**

The sponsors of the LTC asked the Demographic Statistical Methods Division (DSMD) to complete a sample design for each survey, and requested that the designs satisfy certain requirements. However, we found that it was desirable to make some adjustments to these initial requirements.

### **A. Initial Requirements for Blood Sample Survey**

The sponsors requested that *all* LTC respondents in a *sample* of LTC Primary Sampling Units (PSUs) between the ages of 65 and 89 who completed a detailed LTC interview be selected for the BSS. The seventeen largest LTC PSUs (as measured by total 1998 population of persons 65 or older)<sup>1</sup> are to be self-representing (SR); that is, they will be selected with certainty. We refer to the remaining 156 LTC PSUs as non-SR, or NSR. We want to select enough of the NSR PSUs to satisfy the sample requirements shown in Table 1. When the number of NSR PSUs is determined, we will form that number of NSR strata, and select one PSU from each stratum. The stratification will be based on the total 1998 population of persons 65 or older in each PSU.

### **B. Initial Requirements for Buccal Swab Survey**

For the BUCSS, the sponsors requested that a *sample* of LTC respondents ages 80 and older who completed a detailed interview be selected from *all* of the LTC PSUs. The sample size requirements are shown in Table 2.

---

<sup>1</sup> Any reference to the size of a PSU in this memo should be interpreted as the 1998 population of persons 65 or older.

**Table 1: Initial BSS Sample Size Requirements by Age Group**

Age Group	Completed Sample Size Desired	Expected Cooperation Rate	Designated Sample Size
65-69	500	70%	715
70-74	500	70%	715
75-79	350	70%	500
80-84	325	70%	464
85-89	325	70%	464
Total	2,000		2,858

**Table 2: Initial BUCSS Sample Size Requirements by Age Group**

Age Group	Completed Sample Size Desired	Expected Cooperation Rate	Designated Sample Size
80-84	325	90%	361
85-89	325	90%	361
90-94	325	90%	361
95-99	325	90%	361
100+	200	90%	223
Total	1,500		1,667

**C. Adjustments Due to Overlap Between BSS and BUCSS Samples**

Because both the BSS and BUCSS are sampling persons age 80-89, any persons in this age group selected for the BUCSS will also fall into the BSS sample if they are in a PSU selected for the BSS – given the initial requirements described above. In particular, this is true for the seventeen PSUs which are SR for the BSS. Because of concerns about respondent burden, DSMD and the sponsors agreed we should adjust the initial requirements to ensure that no person is selected for both surveys.

We adjusted the requirements in the BSS SR PSUs so that only 50% of LTC respondents age 80-89 are selected for BSS. The BUCSS will draw its sample for this age group from the remaining 50% in these seventeen PSUs.

For the selected BSS NSR PSUs, the BSS will still take 100% of all persons age 65-89. To compensate for the age 80-89 sample in the selected BSS NSR PSUs which is not available for the BUCSS, we will apply weighting factors to persons in that age group who are selected for the BUCSS in the BSS NSR PSUs which are not selected for the BSS. The weighting factors will be determined by stratum. See the section on weighting factors (Section VI ) for details.

***Important Note:*** It is possible that when we form the NSR strata for the BSS, some strata may contain only one PSU. In that event, those PSUs will be considered SR, and the BSS and the BUCSS will split the age 80-89 LTC respondents, as described above for the seventeen largest LTC PSUs.

#### **D. Adjustments to Age Group Sample Size Requirements**

Because of the reduction in the amount of sample available to each survey caused by the adjustments described above, it became evident that it would not be possible to meet the sample size requirements in some age groups for the BUCSS. Also, in order to meet the sample size requirements in some age groups for the BSS, it would be necessary to select many more NSR PSUs than originally expected. This would result in the total sample size (all age groups) for the BSS being much larger than 2,000. Since it is possible to meet the overall sample requirements by reducing the requirements in some age groups and increasing the requirements in others, the sponsors agreed to the following changes:

For the BSS, we will determine the number of NSR PSUs required based on the total sample size, rather than attempting to satisfy the requirement in each age group. This will have the effect of increasing the sample size in some age groups and decreasing it in others.

For the BUCSS, we will re-allocate the requirements between age groups so that the total requirement can be satisfied. The re-allocation will be based on the number of respondents available for BUCSS in each age group after the BSS PSUs have been selected. In those age groups where it is not possible to satisfy the initial requirements listed in Table 2, we will select all the respondents that are available.

### III. Stratification and Selection of NSR PSUs for Blood Sample Survey

#### A. Determining the Number of NSR Strata

To determine the number of BSS NSR strata needed, we first obtained counts by age group of the number of respondents who completed LTC detailed interviews in each LTC PSU. The total number of such respondents age 65- 89 in the seventeen BSS SR PSUs is 1,409. Of these, 711 are age 80-89. Taking away 50% (355) of the 80-89 group, there are 1,054 remaining for the BSS. The total designated sample size for the BSS is 2,858 (from Table 1.) Therefore, we need enough NSR PSUs to make up the difference of  $2,858 - 1,054 = 1,804$ .

In the 156 BSS NSR PSUs, there are 3,960 available respondents age 65-89. Let  $N$  be the number of NSR strata. Then (assuming a reasonably even distribution of respondents among NSR PSUs with similar size populations of 65+) we need the smallest  $N$  such that

$$(N \div 156) \times 3,960 \geq 1,804$$

This is satisfied by  $N = 72$ . Therefore we will form 72 BSS NSR strata.

#### B. Allocating NSR Strata Between the Four Regions

There are four regions defined for the LTC: the Northeast, North Central, South, and West. We determine how many of the 71 BSS NSR strata will be in each region based on the proportion of the total available respondents in NSR PSUs in each region. The results are shown in Table 3.

Note that in the West region, there are two strata which have only one PSU. Therefore, we will convert the two largest BSS NSR PSUs in that region to SR PSUs. This means that 50% of the available respondents age 80-89 will be reserved for the BUCSS sampling in these two PSUs.

#### C. Final Stratification and Selection

In each of the four regions, we sort the BSS NSR PSUs in descending order by size. (The two PSUs in the West which were converted to SR are not included.) All strata will have either two or three NSR PSUs. We form two-PSU strata by grouping every two PSUs, starting with the largest PSUs in each region, until we have formed the number of two-PSU strata indicated for that region in Table 3. If there are any PSUs remaining, they are grouped into three-PSU strata in the same manner. The stratification results are shown in Appendix A.

**Table 3: Allocation of NSR Strata Between Regions**

<b>Region</b>	<b>Number of completed LTC Interviews in BSS NSR PSUs</b>	<b>Proportion <math>P</math> in Region</b>	<b>Number of BSS NSR Strata Allocated (<math>P \times 72</math>)</b>	<b>Total Number of BSS NSR PSUs in Region</b>	<b>Number of PSUs per Stratum</b>
North-east	453	0.11	8	22	2 strata of 2 PSUs, 6 strata of 3 PSUs
North Central	1,094	0.28	20	46	14 strata of 2 PSUs, 6 strata of 3 PSUs
South	1,698	0.43	31	64	29 strata of 2 PSUs, 2 strata of 3 PSUs
West	715	0.18	13	24	2 strata of 1 PSU, 11 strata of 2 PSUs
<b>Total</b>	<b>3,960</b>	<b>1.00</b>	<b>72</b>	<b>156</b>	<b>2 strata of 1 PSU, 53 strata of 2 PSUs, 16 strata of 3 PSUs</b>

Within each strata, we assign each PSU a probability of selection proportional to the number of completed LTC interviews in that PSU, and calculate cumulative probabilities for the PSUs in order of size. We then generate a random number between zero and one for each strata. The PSU selected for each stratum is the one with the smallest cumulative probability larger than the random number. For example, suppose that in a certain stratum there are PSUs A, B, and C. Suppose also that the numbers of LTC interviews in those PSUs are 36, 34, and 30, respectively. Then their cumulative probabilities of selection, respectively, will be 0.36, 0.70, and 1.00. If the random number generated is  $R = 0.5234$ , then PSU B will be selected since 0.70 is the smallest cumulative probability larger than  $R$ . The selection results are also shown in Appendix A.

Table 4 summarizes the sample sizes and expected number of interviews by age group for the BSS. Note that the total number of completed interviews expected is very close to 2,000, although the initial age group requirements are not satisfied.

**Table 4: BSS Sample Sizes and Expected Number of Interview**

<b>Age Group</b>	<b>BSS Sample Size</b>	<b>Expected Cooperation Rate</b>	<b>Expected Number of Complete BSS Interviews</b>
65-69	506	70%	354
70-74	454	70%	318
75-79	740	70%	518
80-84	855	70%	599
85-89	418	70%	293
<b>Total</b>	<b>2,973</b>		<b>2,081</b>

#### **IV. Creating Sample File for Blood Sample Survey**

##### **A. Selection of 80-89 Sample in BSS SR PSUs**

Appendix B lists the seventeen BSS SR PSUs, plus the two NSR PSUs which were converted to SR. Create a file, with all of the cases in the age group 80-89 which are in any of these nineteen PSUs. Call this file the BSS SR universe file. Also create a file to hold all the cases selected for the BSS sample; call this the BSS sample file.

Sort the BSS SR universe file by LTC PSU, SMSA status, age, race, sex, and control number. Select every other record and place the selected records in the BSS sample file. Reserve the remaining cases for BUCSS sampling (section V.)

##### **B. Selection of 65-79 cases in BSS SR PSUs**

Place all of the cases in the BSS SR PSUs, age 65-79, in the BSS sample file.

##### **C. Selection of the Sample in the BSS NSR PSUs**

Place all of the cases age 65-89 from the selected BSS NSR PSUs (except for the two which were converted to SR) into the BSS sample file. The selected BSS NSR PSUs are shown in Appendix A.

Check that the age group distribution in the final BSS sample file matches the counts in the second column of Table 4.



**V. Selection of Buccal Swab Survey Sample**

**A. Sample Available for BUCSS**

Once the BSS sample has been selected, all of the remaining completed LTC interviews with respondents age 80 and older are eligible to be selected for the BUCSS. The second column in Table 5 shows how many cases are available for the BUCSS in each age group.

**B. Reallocation of BUCSS Sample Size Requirements**

We wish to reallocate the BUCSS age group sample size requirements shown in Table 2 so that the total requirement can be satisfied. The third column in Table 5 shows the original designated sample size in each age group, and the fourth column shows how many more cases are needed. We decrease the designated sample size for the 95-99 and 100+ age groups to the number of cases actually available in those groups. Then we allocate the 174 cases needed to satisfy the total requirement among the other three age groups based on the amount of extra sample they have remaining, shown in columns five and six of Table 5. The last column in Table 5 shows the amount to increase the designated sample size in each of the age groups. Table 6 shows the new designated sample size and the number of completed interviews expected for each age group.

**Table 5: Sample Available for the BUCSS**

<b>Age Group</b>	<b>Number of Cases Available</b>	<b>Original Designated Sample Size</b>	<b>Number of Cases Short</b>	<b>Remainder Available</b>	<b>Age Group Percent of Total Remainder</b>	<b>Number to Add to Age Group Designated Sample Size</b>
80-84	910	361	0	549	79%	139
85-89	443	361	0	82	12%	21
90-94	423	361	0	62	9%	16
95-99	309	361	52	0	0%	0
100+	101	223	122	0	0%	0
<b>Total</b>	<b>2,186</b>	<b>1,667</b>	<b>174</b>	<b>693</b>	<b>100%</b>	<b>174</b>

Table 6: Final Designated Sample Sizes and Expected Completions for the BUCSS

Age Group	Final Designated Sample Size	Expected Cooperation Rate	Expected Number of Completed Interviews
80-84	499	90%	449
85-89	381	90%	343
90-94	377	90%	339
95-99	309	90%	278
100+	101	90%	91
<b>Total</b>	<b>1,667</b>		<b>1,500</b>

### C. Selecting the BUCSS Sample

Select the BUCSS sample within each age group from the 80+ LTC respondents which remain after the BSS sample is selected. Create separate files for each age group: 80-84, 85-89, 90-94, 95-99, and 100+. Call these the BUCSS universe files. Also create a file to hold the selected BUCSS sample; call this the BUCSS sample file.

Within each universe file, sort by LTC PSU, SMSA status, age, race, sex, and control number. For each age group universe file, do the following:

1. Let  $N$  be the number of records in the file.
2. Let  $k$  be the designated sample size.
3. Let  $SI = N \div k$  be the sampling interval.
4. Number the cases in the sorted file for the age group consecutively from 1 to  $N$ .
5. Generate a random number  $X$  between 0 and 1, and let  $RS = X \times SI$  be the random start.
6. Generate the sequence  $\{c_1, c_2, \dots, c_k\}$ , where

$$c_i = RS + [SI \times (i-1)] , \quad i = 1, 2, 3, \dots, k$$

7. For each  $i$  between 1 and  $k$ , let  $h_i = \lceil c_i \rceil$ ; that is, let  $h_i$  be the smallest positive integer greater than or equal to  $c_i$ . (The function  $f(x) = \lceil x \rceil$  is called the *ceiling function*.)

8. Select the cases in the numbered universe file which have numbers corresponding to the numbers in the sequence  $\{h_1, h_2, h_3, \dots, h_k\}$  and place them in the BUCSS sample file.
9. Repeat this process for each of the five age group universe files.

Finally, check to make sure that the age group distribution in the BUCSS sample file matches the second column in Table 6.

## VI. Weighting Factors for Blood and Buccal Swab Surveys

Since both the BSS and BUCSS samples are sub-samples of the LTC Survey, we can calculate "base" weights for the BSS and BUCSS by applying sub-sampling factors to the LTC final weight for each case. (We will not discuss noninterview adjustment here.)

Table 7 indicates how the factors may be calculated. The variables *SI* and *POS* in the table are defined as follows:

*SI* = BUCSS sampling interval for each age group (see V.C.3.)

*POS* = probability of selection within a BSS stratum for the selected BSS PSU

Appendix C lists all of the factors by PSU and age group for the BSS and BUCSS.

**Table 7: Calculation of Weighting Factors for the BSS and BUCSS**

Age Group	SR		NSR	
	BSS	BUCSS	BSS	BUCSS
65-79	1	–	$\frac{1}{POS}$	–
80-89	2	$2 \times SI$	$\frac{1}{POS}$	$\frac{1}{(1-POS)} \times SI$
90+	–	<i>SI</i>	–	<i>SI</i>

## VII. Verification of Sampling for Blood and Buccal Swab Surveys

For verification of the sampling, please provide DSMD with the following files (preferably as SAS data sets):

1. BSS Sample File [Section IV]
2. BUCSS Universe File for 80-84 Age Group [Section V.C.]
3. BUCSS Universe File for 85-89 Age Group [Section V.C.]
4. BUCSS Universe File for 90-94 Age Group [Section V.C.]
5. BUCSS Universe File for 95-99 Age Group [Section V.C.]
6. BUCSS Universe File for 100 + Age Group [Section V.C.]
7. BUCSS Sample File [Section V.C.]

Each file should include the following variables (the names should match variables on the LTC weighting file):

1. CTRLNUM
2. SMSA
3. LTC\_PSU
4. I\_AGE
5. RACE
6. SEX

Also, for the BUCSS Universe Files, please provide the variables  $N$ ,  $k$ ,  $SI$ , and  $RS$  used in sampling from that age group. These variables are defined in Section V.C.

## VIII. References

- [1] Memorandum for Corder (Duke University) from Murphy, "Sample Design Plans for Blood Sample Survey and Buccal Swab Survey," February 17, 2000
- [2] Memo for Bowie from Tupek, "1999 Long-Term Care (LTC) Survey Weighting Specifications for Cross-Sectional Estimates," *[still in draft form as of May 26, 2000]*
- [3] Memorandum for Bowie from Tupek, "Sampling Specifications for the Long-Term Care Survey (LTC): Longitudinal Portion and Longitudinal Portion of the Healthy Supplement and 95+ Supplement," May 18, 1999
- [4] Memorandum for Bowie from Tupek, "Sampling Specifications for the Long-Term Care Survey: the Aged-In Cohort, the 95+ Supplement, and the Aged-In Portion of the Healthy Supplement," May 18, 1999

## Stratification and Selection of NSR PSUs for the Blood Sample Survey (BSS)

## Northeast Region

BSS stratum	LTC PSU	1998 Population Age 65+	Number of Completed LTC Interviews Age 65-89	Prob. Of Selection (POS)	Cum. POS	Random Number	Selection
NE-1	107	183,701	33	0.6346	0.6346	0.9235	X
	109	154,328	19	0.3654	1.0000		
NE-2	105	117,360	11	0.3333	0.3333	0.6957	X
	154	111,938	22	0.6667	1.0000		
NE-3	106	111,232	15	0.2083	0.2083	0.2475	X
	121	102,784	21	0.2917	0.5000		
NE-4	129	86,225	36	0.5000	1.0000	0.9172	X
	125	82,209	21	0.3559	0.3559		
	131	75,984	21	0.3559	0.7119		
	123	71,860	17	0.2881	1.0000		
NE-5	122	62,897	21	0.3621	0.3621	0.0603	X
	130	48,911	28	0.4828	0.8448		
NE-6	124	48,824	9	0.1552	1.0000	0.6757	X
	127	45,290	27	0.5000	0.5000		
	126	43,518	21	0.3889	0.8889		
	128	42,365	6	0.1111	1.0000		
NE-7	152	33,643	31	0.4429	0.4429	0.0381	X
	153	22,397	19	0.2714	0.7143		
NE-8	151	18,442	20	0.2857	1.0000	0.4923	X
	156	13,341	14	0.2545	0.2545		
	157	8,004	22	0.4000	0.6545		
	155	4,001	19	0.3455	1.0000		