UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

MAY 1 8 1999

MEMORANDUM FOR Chester E. Bowie
Chief, Demographic Surveys Division

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

Subject: Sampling Specifications for the 1999 Long-Term Care Survey: the Aged-In Cohort, the 95+ Supplement, and the Aged-In Portion of the Healthy Supplement

I. SUMMARY OF SAMPLE COHORTS AND SUPPLEMENTS FOR THE 1999 LONG-TERM CARE SURVEY (LTC)

The 1999 LTC sample consists of six cohorts and five supplements. Five of the six cohorts consist of persons eligible for the 1994 LTC. They comprise the longitudinal portion of the 1999 LTC. The longitudinal portion includes the following 5 cohorts:

Cohort 1 -- 3,557 persons with a community interview in 1994;
Cohort 4 -- 334 persons with an institutional interview in 1994;
Cohort 5 -- a sample of 5,000 persons from persons unimpaired in the 1994 survey;
Cohort 7 -- all remaining persons unimpaired in 1994 who are 75 years and older on or before March 31, 1999 (expected sample size is 3,284); and
Cohort 8 -- 1,600 persons previously selected for the 1989 survey who were unimpaired in 1989 but were not selected for the 1994 survey.

The longitudinal portion of the 1999 LTC can be used for both cross-sectional and longitudinal estimates.

The sixth cohort can be used for cross-sectional estimates only. This cohort is usually referred to as the aged-in cohort. The aged-in cohort consists of the following:

Cohort 6 -- a sample of 5,500 Medicare persons who reached their 65th birthday between April 1, 1994 and March 31, 1999, inclusive. (In other words, people who turned 65 since the 1994 LTC.)

The sample for the three 1999 LTC supplements consist of the following:

- Cohort 9 -- a sample of 600 persons who reached their 95th birthday between

April 1,1994 and March 31, 1999 (the 95+ supplement)

- a sample of 1,545 unimpaired persons aged 65 and over (the healthy supplement). Approximately 1,262 of these people continue from the 1994 survey sample and about 283 will be newly selected from cohorts 6 and 8 above

- all known secondary care givers to any sample person (the Care giver supplement) [Note that since interviewers will attempt to contact *every* Care giver discovered during the survey, no sampling is necessary; so there are no procedures involving the Care giver supplement discussed in this document.]

This memorandum gives instructions for selecting the aged-in cohort, the 95+ supplement, and the aged-in cohort of the healthy supplement. We will administer an abbreviated community questionnaire to aged-in persons designated for the healthy supplement if they are determined to be healthy during the screening interview. If they are determined to be impaired during the screening interview, they will get the entire questionnaire.

A second memorandum specifies the selection of the longitudinal cohorts and assignment of persons from the longitudinal cohorts to be in the healthy supplement and the 95+ supplement.

II. OVERVIEW OF SELECTING THE AGED-IN COHORT, THE 95+ SUPPLEMENT, AND THE AGED-IN PORTION OF THE HEALTHY SUPPLEMENT

Specifications for extracting the 1999 Medicare subfile are given in a memorandum, from Barbara Walter to Chester E. Bowie, dated May 19, 1998, titled "Specifications to Extract the 1999 Medicare Subfile for the 1999 Long-Term Care Survey (LTC)." The DSD will divide the 1999 Medicare subfile into two files. The first file will consist of persons who become 65 years old between April 1, 1994 and March 31, 1999, inclusive. This file will be referred to as the "first 1999 universe subfile." The second file will consist of persons who become 95 years old before April 1, 1999 and will be referred to as the "second 1999 universe subfile."

DSD will use the first 1999 universe subfile for selecting the aged-in cohort and assigning the aged-in portion of the healthy supplement. First DSD will sort the file and obtain various counts. Using these counts, DSD will select the aged-in cohort from the first 1999 universe subfile. After selecting the aged-in cohort, DSD will systematically select a sample for the healthy supplement from the persons selected for the aged-in cohort. These persons will be designated with an H in their control number. They will be given an abbreviated community questionnaire provided they are determined to be

healthy on the screener interview. Otherwise they will receive a full interview.

The 95+ supplement will be selected from part of the second 1999 universe subfile. First DSD will sort the file and obtain various counts. Then DSD will match the second 1999 universe subfile to the 1999 LTC longitudinal file to determine which persons aged 95 and over are already in the longitudinal sample. DSD will delete the longitudinal persons aged 95 and over from the second 1999 universe subfile. DSD will verify that nonmatched persons on the longitudinal file are recently deceased or have moved out of LTC PSUs. The remaining file will be referred to as the "95+ eligible subfile." DSD will obtain various counts for the 95+ eligible subfile. Using these counts, they will then select the sample for the 95+ supplement. Anyone who was in the 95+ supplement in 1994 will remain in the 95+ supplement in 1999. After all the samples are selected, DSD will form a sample file with the items specified in Attachment B and the codes and parameters specified in section X.

III. LTC PSU SELECTION

The LTC sample is a subsample of the 1970 current survey A-design sample. An A-design PSU consists of a county or a group of contiguous counties. All counties in the entire United States are grouped into 376 A-design strata. One hundred fifty-six of the strata consist of only one PSU; these are called self-representing (SR) PSUs. The remaining 220 strata are formed from counties with similar total 1970 census population and characteristics such as SMSA, percent of population urban, percent of population with race as nonwhite, number of farms, and per capita retail sales. A PSU is then selected from each A-design stratum with probability proportional to its 1970 population. These 220 sample PSUs are called non-self-representing (NSR) PSUs. More detailed sampling procedures for the A-design sample are given in "Technical Paper 40, The Current Population Survey--Design and Methodology."

For LTC, A-design strata are collapsed to form 173 LTC strata. Thirty-nine of the LTC strata consist of only one PSU. These PSUs are automatically included in the LTC sample and are called LTC self-representing (SR) PSUs. The remaining 134 LTC strata are formed by combining A-design strata with similar characteristics. The combining of A-design strata into LTC strata is based on the proportion of the population 65 years and older and enrolled in Medicare (referred to below as Medicare population) as estimated for the A-design sample PSU in each A-design stratum.

In each of the remaining 134 LTC strata of approximately equal Medicare population, one A-design stratum is selected with probability proportional to its estimated 1978 Medicare population. The A-design sample PSU from that stratum is then selected as an LTC sample PSU. These sample PSUs are called LTC non-self-representing (NSR) PSUs. A list of LTC PSUs (both SR and NSR) appears in Attachment A.

IV. PRELIMINARY OPERATIONS FOR SELECTING THE SAMPLES

A. The Creation of the 1999 Universe Subfiles From the 1999 Medicare Subfile

Prior to sampling, it will be necessary to form strata for the aged-in sample, to assign an LTC PSU number to each record, to assign a current surveys (CS) PSU number to each record, to divide the Medicare subfile into two universe subfiles, and to sort each universe subfile. The final products will be referred to as the first and second 1999 universe subfiles.

1. Form strata.

   a. For the aged-in sample, form strata based on each person's original reason for entitlement (first occurrence) of BENE_ENTLMT_RSN_CD

      (i) Stratum 1 - BENE_ENTLMT_RSN_CD = 0

      This stratum contains all persons originally entitled due to age.

      (ii) Stratum 2 - BENE_ENTLMT_RSN_CD = 1, 2, 3

      This stratum contains all persons originally entitled by disability.

      (iii) If the original reason for entitlement is blank, then delete these records. They are not part of the Medicare universe

   b. For the 95+ supplement, the universe will not be stratified.

      If the original reason for entitlement code is blank, include these records in the universe. Persons 95 and over may not have entered the Medicare system because the system was started in the 1970's after these persons were already 65.

2. Assign the appropriate LTC PSU number to each record. Also assign the appropriate CS PSU number to each record in each subfile. In most cases an LTC PSU consists of only one CS PSU. For those LTC PSU's containing more than one CS PSU, a CS PSU number is listed in Attachment C for each county in each LTC PSU. For counties lying in more than one CS PSU, the lowest CS PSU number of the LTC PSU is listed.

3. Divide the 1999 Medicare subfile into two files:

Refer to the first file as the "first 1999 universe subfile." It will consist of persons who turned 65 between April 1, 1994 and March 31, 1999, inclusive and are alive on March 31, 1999.

Refer to the second file as the "second 1999 universe subfile." It will consist of persons who turned 95 on or before March 31, 1999 and are alive on March 31, 1999.

4. Sort each file in the following manner:

   a. By stratum number (j=1 or 2)

   b. By LTC PSU number within stratum (i=1-173)

   c. By race (white, black, other, unknown) within LTC PSUs within stratum

   d. By age within races within LTC PSUs within stratum

B. The Creation of the 95+ Eligible Subfile from the Second 1999 Universe Subfile

1. Match the second 1999 universe subfile to the 1999 LTC longitudinal file.

2. Delete the persons aged 95 and over who match to the 1999 LTC longitudinal file.

3. Verify with the Health Care Financing Administration that nonmatched persons on the longitudinal file are recently deceased or have moved out of an LTC PSU.

   All persons aged 95 and over on the longitudinal file should be on the second 1999 universe subfile. A discrepancy could occur one of two ways. First, the longitudinal file was updated as of early February 1999. The second 1999 universe was updated as of March 31, 1999. Refer to memo from Tupek to Bowie, dated May 18, 1999, "Specifications for Updating the Longitudinal Portion of the Long-Term Care Survey (LTC) for 1999." Nonmatched longitudinal sample could occur if persons on the longitudinal file have died between early February 1999 and March 31, 1999, inclusive. Second, nonmatched cases could occur if longitudinal persons moved out of an LTC PSU.

4. Refer to the remaining second 1999 universe file as the "95+ eligible subfile" (i.e., these persons are on the second 1999 universe subfile and

not on the 1999 LTC longitudinal file).

C. Obtaining Counts of the 1999 Universe Subfiles and the 95+ Eligible Subfile

Obtain the following counts of persons:

N_{ij1} = the total number of persons in the i^{th} LTC PSU and j^{th} stratum from the first 1999 universe subfile.

$N_{i.1}$ = the total number of persons in the i^{th} LTC PSU from the first 1999 universe subfile.

Note that N_{ij1} for Group A PSUs (non-asterisked in Attachment A) represent 10 percent and N_{ij1} for Group B PSUs (asterisked in Attachment A because they are smaller PSUs) represent 50 percent of all persons on the 1999 Health Insurance Skeleton Eligibility Write-Off file who become 65 years old on or after April 1, 1994 and are alive as of March 31, 1999 in the i^{th} LTC PSU and j^{th} stratum. The same is true for $N_{i.1}$ except that $N_{i.1}$ includes all stratum in the i^{th} PSU.

$N_{i.2}$ = the total number of persons in the i^{th} LTC PSU from the second 1999 universe subfile.

$E_{i.2}$ = the total number of persons 95 years of age or older on April 1, 1999 in the i^{th} LTC PSU from the 95+ eligible subfile.

V. PRELIMINARY OPERATIONS FOR SELECTING THE AGED-IN COHORT

A. Sample Size and Sampling Intervals

The national sampling interval (si) for the 1999 LTC aged-in cohort should be calculated using N_{ij1} and π_i. π_i is the probability of selection which remains the same as the 1982 LTC design. (The π_i's are listed in Attachment A.) The si should result in a designated sample of approximately 5,500 persons.

1. The national sampling interval for selecting the aged-in cohort,

$$si = \frac{\sum\sum f_i N_{ij1}/\pi_i}{5500}$$

$f_i = 10$ for each Group A PSU (non-asterisked PSUs in Attachment A).
$f_i = 2$ for each Group B PSU (asterisked PSUs in Attachment A).

2. For each LTC PSU within each stratum within the first 1999 universe subfile, compute

$\hat{n}_{ij1}$ = the expected sample size for the i^{th} LTC PSU in the j^{th} stratum for the first 1999 universe subfile

$$= \frac{10N_{ij1}}{\pi_i s_i} \text{ rounded to 4 decimal places, for Group A PSUs}$$

$$= \frac{2N_{ij1}}{\pi_i s_i} \text{ rounded to 4 decimal places, for Group B PSUs.}$$

For several LTC PSUs $\hat{n}_{ij1}$ may be greater than N_{ij1}. If this occurs do the following:

1. Assign $w_{ij1} = \hat{n}_{ij1}$
2. Let $\hat{n}_{ij1} = N_{ij1}$.

3. For each stratum within the first 1999 universe subfile, compute

$$\hat{n}_{.j1} = \sum_{i}^{173} \hat{n}_{ij1} .$$

4. For each LTC PSU within the first 1999 universe subfile, compute

$$\hat{n}_{i.1} = \sum_{j}^{2} \hat{n}_{ij1} .$$

5. Also compute

$$\hat{n} = \sum_{j}^{2} \hat{n}_{.j1} = \sum_{i}^{173} \hat{n}_{i.1}$$

(perform the calculation both ways as a check).

B. Computing Take Everys and Random Starts for the Aged-in Cohort

A take every ($TE_{i.1}$) is required for each LTC PSU and a random start (RS_{ij1}) is required for each LTC PSU and stratum. These parameters are used in the sample selection procedure outlined below.

Note that the aged-in cohort will only be selected from the first 1999 universe subfile and the take every does not vary by stratum.

Perform the following operations on the first 1999 universe subfile for each LTC PSU:

1. Compute the take every:

$$TE_{i.1} = \frac{\pi_i si}{10}$$, rounded to 4 decimal places, for Group A PSUs.

$$TE_{i.1} = \frac{\pi_i si}{2}$$, rounded to 4 decimal places, for Group B PSUs.

If $TE_{i.1} < 1.0000$ for any LTC PSU, provide the LTC PSU number and the value of $Te_{i.1}$ to DSMD, then let $TE_{i.1}=1$.

2. For each stratum, determine the random start associated with the j^{th} stratum and first LTC PSU,

$RS_{1j1} = 1$ if $TE_{1.1} = 1$, otherwise

$RS_{1j1} = \alpha_j TE_{1.1}$, rounded to 4 decimal places,

where α_j is a 4 decimal place random number such that $0 < \alpha_j \leq 1.0000$. Thus, it should be true that $0 < RS_{1j1} \leq TE_{1.1}$. The $\alpha_{j's}$ can be generated by the computer using a random number generator. RS_{ij1} is only needed here for the first LTC PSU in each stratum; all other random starts will be generated by an algorithm using parameters from the sampling of the previous LTC PSU and stratum. (Described in section VI.A.1c.) All $TE_{i.1}$'s and RS_{ij1}'s should be saved for output later.

VI. SELECTING THE AGED-IN COHORT OF MEDICARE ENROLLEES

A. Sample Selection

Select the aged-in cohort of 1999 Medicare enrollees from the first 1999 universe subfile from each LTC PSU within each stratum in the order in which it is sorted as specified below.

1. Sampling the First 1999 Universe Subfile for the Aged-in Cohort

Sample the first 1999 universe subfile in the sort order described in Section IV.A.4. Sample the first LTC PSU in stratum 1 using $TE_{1.1}$, followed by the second LTC PSU in stratum 1 using $TE_{2.1}$, etc., until the stratum has been entirely sampled. Sample stratum 2 in the same way as stratum 1.

Each LTC PSU has one take every $TE_{i.1}$ which is to be used in sampling both strata for that LTC PSU. However, within a stratum, determine the random start RS_{ij1} used in sampling a LTC PSU from parameters associated with sampling the previous LTC PSU, continuing in this fashion until the stratum is entirely sampled. Generate the first RS for each stratum as described in Section V.B.2.

a. Applying RS_{ij1} and $TE_{i.1}$

Select the sample of persons from the i^{th} LTC PSU, and j^{th} stratum, in the following manner:

Determine the sequence of numbers

RS_{ij1}, $RS_{ij1} + TE_{i.1}$, $RS_{ij1} + 2TE_{i.1}$,....,

until the absolute difference between N_{ij1} and the last member of the sequence is less than $TE_{i.1}$.

Next, round each member of the sequence up to the next integer (e.g., 6.0487 is rounded to 7). A number such as 5.0000 is rounded to 5. Note: It is important to round UP so that the first and last records don't have a higher/lower probability of selection.

Consider the persons in the i^{th} LTC PSU, and the j^{th} stratum as numbered consecutively from 1 to N_{ij1}. Then those persons with numbers corresponding to those in the above rounded sequence are

the sample persons from that stratum and LTC PSU.

Note: if $TE_{i.1} = 1$ then the sequence of numbers for the i^{th} PSU and the j^{th} stratum will be 1, 2, 3...N_{ij1}.

b. Verifying the Sample Selection

Within a stratum the accuracy of the sample selection in one LTC PSU should be checked before proceeding to the next LTC PSU. Sampling should not continue if in any cases (i) and (ii.) below do not hold.

(i) n_{ij1}, the actual sample size obtained from sampling the i^{th} LTC PSU, and j^{th} stratum should be within one of the expected sample size $\hat{n}_{ij1}$. That is, it should be true that

$| n_{ij1} - \hat{n}_{ij1} | \leq 1.0000$ unless this is a PSU where

$\hat{n}_{ij1} > N_{ij1}$. If $TE_{i.1} = 1.0000$ then n_{ij1} should equal N_{ij1}.

Notify DSMD if there is a discrepancy between the actual and expected sample sizes unless this is a PSU where $TE_{i.1}$ was originally less than 1.

(ii) If $N_{ij1} > RS_{ij1}$, then define LO_{ij1}, the leftover from sampling the i^{th} LTC PSU, and j^{th} stratum, as N_{ij1} minus the last member of the unrounded sequence determined above. The expected leftover from sampling the i^{th} LTC PSU, and j^{th} stratum, is defined as:

$$EXPLO_{ij1} = N_{ij1} - \left[RS_{ij1} + \left(n_{ij1} - 1\right) TE_{i.1}\right]$$

It should always be true that $LO_{ij1} = EXPLO_{ij1}$. Notify DSMD in the case that $LO_{ij1} \neq EXPLO_{ij1}$.

If $N_{ij1} < RS_{ij1}$, then $LO_{ij1} = N_{ij1}$.

c. Determine RS_{ij1} for successive LTC PSUs within a stratum

(i) After sampling the i^{th} LTC PSU in the j^{th} stratum, determine

the next random start as follows (assuming the i^{th} LTC PSU is not the last to be sampled in the stratum):

If the i^{th} LTC PSU is not the last to be sampled in the j^{th} stratum, then

$$RS_{(i+1)j1} = \left(\frac{TE_{i.1} - LO_{ij1}}{TE_{i.1}} \right) TE_{(i+1).1} ,$$

rounded to 4 decimal places

(ii) If the i^{th} LTC PSU is the last to be sampled in stratum 1, then it is not necessary to save LO_{i11} and RS_{i11}.

d. Example:

Suppose the first LTC PSU in stratum j for the first 1999 universe subfile is a Group A LTC PSU and has $\pi_1 = .051123$ and $N_{1j1} = 20$

and si = 730.590873, $\hat{n}_{1j1} = \frac{10N_{1j1}}{\pi_1 si} = 5.3548$

$$TE_{1.1} = \frac{\pi_1 si}{10} = 3.7350$$

Let $\alpha_j = .3476$, then $RS_{1j1} = \alpha_j TE_{1.1} = 1.2983$.
The sequence of numbers is calculated as 1.2983, 5.0333, 8.7683, 12.5033, 16.2383, 19.9733 and the rounded sequence is 2, 6, 9, 13, 17, 20 thus the 2nd, 6th, 9th, 13th, 17th, and 20th persons in the first LTC PSU in stratum j on the sorted universe subfile are sample persons.

It can be seen that $n_{1j1} = 6$ is within one of $\hat{n}_{1j1} = 5.3548$. The expected leftover is

$$EXPLO_{1j1} = N_{1j1} - [RS_{1j1} + (n_{1j1} - 1)TE_{1.1}]$$

$$= 20 - [1.2983 + (5)(3.7350)]$$

$= 20 - 19.9733$

$= .0267$

and the actual leftover is

$LO_{1j1} = 20 - 19.9733 = .0267.$

So $LO_{1j1} = EXPLO_{1j1}$. It can be seen that $LO_{1j1} < TE_{1.1}$. The take every for sampling the second LTC PSU, j^{th} stratum is $TE_{2.1} = \pi_2 si$, and the random start is

$$RS_{2j1} = \left(\frac{TE_{1.1} - LO_{1j1}}{TE_{1.1}} \right) TE_{2.1}$$

$$= \left(\frac{3.7350 - .0267}{3.7350} \right) TE_{2.1}$$

$=(.9929)TE_{2.1}$, rounded to 4 decimal places

e. Sample Counts

After sampling the first 1999 universe subfile, determine

n_{ij1} = number of persons sampled from the i^{th} LTC PSU and j^{th} stratum

$n_{i.1}$ = number of persons sampled from the i^{th} LTC PSU

$$= \sum_{j}^{2} n_{ij1} \quad .$$

B. Sample File

Following the sample selection, create a sample file consisting of the entire compressed record (see Attachment B), plus the LTC PSU number and all sample identifiers and parameters listed in section X for each person selected for the aged-in cohort from the first 1999 universe subfile.

VII. DESIGNATING WHICH AGED-IN PERSONS ARE ELIGIBLE FOR THE HEALTHY SUPPLEMENT

A. Overview

A sample of persons from the aged-in cohort need to be selected for the healthy supplement. These persons will screen in as healthy and will be interviewed with the Community Questionnaire at the time of their LTC interview. This section describes the systematic selection of sample persons from the aged-in cohort. Section XIII describes how an "H" will be assigned to the control number for persons designated for the healthy supplement.

B. Sorting the Aged-in Cohort

Before selecting the healthy supplement from the aged-in cohort, sort the records for the aged-in cohort in the following order:

1. By stratum number

2. By LTC PSU number within stratum number

3. By race (white, black, other, unknown) within LTC PSUs within stratum

4. By age within races within LTC PSUs within stratum

C. Random Start and Sampling Interval

1. The sampling interval (SI) for selecting the healthy supplement is

(1,600 (persons from cohort 8) + 5,500 (aged-in)) ÷ 283 = 7100/283 = 25.0883

2. Generate U, a random number between 0 and 1. Then calculate the random start: RS = U * SI.

D. Applying the Random Start and Sampling Interval

Select the sample of persons for the healthy supplement as follows:

1. First determine the sequence of numbers

   RS, RS + SI, RS + 2*SI, RS + 3*SI, ...,until the absolute difference between the number of persons selected for the aged-in cohort (n) and the last member of the sequence is less than SI.

2. Next round each number of the sequence up to the next integer (e.g., 6.0487 is rounded to 7). Note a number such as 5.0000 is rounded to 5.

3. Consider the aged-in cohort as numbered consecutively from 1 to n.

Those persons with numbers corresponding to those in the above rounded sequence are the persons from the aged-in cohort that are in the healthy supplement.

E. Designating Persons From the Aged-in Cohort to Receive the Healthy Supplement if They Screen In as Healthy

The records of persons selected for the aged-in cohort and the healthy supplement will be designated with an H in their control numbers according to the specifications given in section XIII.

VIII. PRELIMINARY OPERATIONS FOR SELECTING THE 95+ SUPPLEMENT

A. Sample Size and Sampling Intervals

Using the 95+ eligible subfile, the national sampling interval (ri) for the 1999 LTC 95+ supplement should be calculated using $E_{i.2}$ and π_i. π_i is the probability of selection which remains the same as the 1982 LTC design. (The π_i's are listed in Attachment A.) The sampling interval should result in a designated sample of approximately 600 persons 95 years of age or older. Note, this 600 is in addition to the 95+ from 1994 which will remain in the supplement in 1999. We estimate 32 persons from cohort 1 of the 1994 95+ supplement will be in 1999; 15 persons from cohort 4 of the 1994 95+ supplement will be in 1999; and 131 persons from cohorts 5 and 7 of the 1994 95+ supplement will be in 1999.

1. The national sampling interval for selecting the 95+ supplement is:

$$ri = \frac{\sum E_{i.2}/\pi_i}{600} .$$

2. For each LTC PSU within the 95+ eligible subfile, compute

$\hat{e}_{i.2}$ = the expected sample size for the i^{th} LTC PSU

$$= \frac{E_{i.2}}{\pi_i ri}$$ rounded to 4 decimal places.

For several LTC PSUs $\hat{e}_{i.2}$ may be greater than $E_{i.2}$. If this occurs do

the following:

1. Assign $w_{i.2} = \hat{e}_{i.2}$

2. Let $\hat{e}_{i.2} = E_{i.2}$.

3. Compute

$$\hat{e}_{..2} = \sum_i \hat{e}_{i.2}$$

B. Computing Take Everys and Random Starts for the 95+ Supplemental Sample

A take every ($TE_{i.2}$) and a random start ($RS_{i.2}$) are required for each LTC PSU. These parameters are used in the sample selection procedure outlined below.

Perform the following operations on the 95+ eligible subfile for each LTC PSU:

1. Compute the take every:

$TE_{i.2} = \pi_i ri$, rounded to 4 decimal places.

If $TE_{i.2} < 1.0000$ for any LTC PSU, print out the LTC PSU number and the value of $TE_{i.2}$, then let $TE_{i.2}=1$.

2. Determine the random start associated with the first LTC PSU,

$RS_{1.2} = 1$ if $TE_{1.2} = 1$, otherwise

$RS_{1.2} = \alpha_j TE_{1.2}$, rounded to 4 decimal places,

where α_j is a 4 decimal place random number such that $0 < \alpha_j \leq 1.0000$. Thus it should be true that $0 < RS_{1.2} \leq TE_{1.2}$. The $\alpha_{j's}$ can be generated by the computer using a random number generator. $RS_{i.2}$ is only needed here for the first LTC PSU; all other random starts will be generated by an algorithm using parameters from the sampling of the previous LTC PSU. All $TE_{i.2}$'s and $RS_{i.2}$'s should be saved for output later.

IX. SELECTING THE 95+ SUPPLEMENT OF MEDICARE ENROLLEES

A. Sample Selection

Select the 95+ supplemental sample of 1999 Medicare enrollees from each LTC PSU in the order in which it is sorted as specified below.

1. Sampling the 95+ Eligible Subfile for the 95+ Supplement

Sample the 95+ eligible subfile in the sort order described in Section IV.A.4. The first LTC PSU in the 95+ eligible subfile will be sampled using $TE_{1.2}$, followed by the second LTC PSU using $TE_{2.2}$, etc.

The random start $RS_{i.2}$ used in sampling a LTC PSU is determined from parameters associated with sampling the previous LTC PSU. The random start is generated as described in Section VIII.B.2.

a. Applying $RS_{i.2}$ and $TE_{i.2}$

Select the sample of persons from the i^{th} LTC PSU in the following manner:

Determine the sequence of numbers

$RS_{i.2}$, $RS_{i.2} + TE_{i.2}$, $RS_{i.2} + 2TE_{i.2}$,....,

until the absolute difference between $E_{i.2}$ and the last member of the sequence is less than $TE_{i.2}$.

Next, round each member of the sequence up to the next integer (e.g., 6.0487 is rounded to 7). Note that a number such as 5.0000 is rounded to 5.

Consider the persons in the i^{th} LTC PSU as numbered consecutively from 1 to $E_{i.2}$. Then those persons with numbers corresponding to those in the above rounded sequence are the sample persons from that LTC PSU.

Note: If $TE_{i.2} = 1$ then the sequence of numbers for the i^{th} PSU will be 1, 2, 3...$E_{i.2}$.

b. Verifying the Sample Selection

The accuracy of the sample selection in one LTC PSU should be checked before proceeding to the next LTC PSU. Sampling should not continue if in some case (i) and (ii) below do not hold.

(i) $e_{i.2}$, the actual sample size obtained from sampling the i^{th} LTC PSU in the 95+ eligible subfile should be within one of the expected sample size $\hat{e}_{i.2}$. That is, it should be true that $| e_{i.2} - \hat{e}_{i.2} | \leq 1.0000$ unless this is a PSU where $\hat{e}_{i.2} > E_{i.2}$. If $TE_{i.2} = 1.0000$ then $e_{i.2}$ should equal $E_{i.2}$.

Notify DSMD if there is a discrepancy between the actual and expected sample sizes, unless this is a PSU where $TE_{i.2}$ was originally less than 1.

(ii) Define $LO_{i.2}$, the leftover from sampling the i^{th} LTC PSU as $E_{i.2}$ minus the last member of the unrounded sequence determined above. The expected leftover from sampling the i^{th} LTC PSU is defined as:

$$EXPLO_{i.2} = E_{i.2} - \left[RS_{i.2} + \left(e_{i.2} - 1\right) TE_{i.2}\right]$$

It should always be true that $LO_{i.2} = EXPLO_{i.2}$. Notify DSMD in the case that $LO_{i.2} \neq EXPLO_{i.2}$.

c. Determine $RS_{i.2}$ for successive LTC PSUs

(i) After sampling the i^{th} LTC PSU, determine the next random start as follows (assuming the i^{th} LTC PSU is not the last to be sampled):

If the i^{th} LTC PSU is not the last to be sampled, then

$$RS_{(i+1).2} = \left(\frac{TE_{i.2} - LO_{i.2}}{TE_{i.2}} \right) TE_{(i+1).2} ,$$

rounded to 4 decimal places

(ii) If the i^{th} LTC PSU is the last to be sampled, then it is not necessary to save $LO_{i.2}$ and $RS_{i.2}$.

d. Example:

Suppose the first LTC PSU for the 95+ eligible subfile has $\pi_1 =$ .051123, $E_{1.2} = 12$, and ri = 60.129, $\hat{e}_{1.2} = \frac{E_{1.2}}{\pi_1 ri} = 3.9037$

$TE_{1.2} = \pi_1 ri = 3.0740$.

Let $\alpha_j = .4221$, then $RS_{1.2} = \alpha_j TE_{1.2} = 1.2975$.

The sequence of numbers is calculated as 1.2975, 4.3715, 7.4455, 10.5915 and the rounded sequence is 2, 5, 8, and 11. Thus the 2nd, 5th, 8th, and 11th persons in the first LTC PSU on the sorted 95-99 eligible subfile are sample persons.

It can be seen that $e_{1.2} = 4$ is within one of $\hat{e}_i.2 = 3.9037$. The expected leftover is

$$EXPLO_{1.2} = E_{1.2} - [RS_{1.2} + (e_{1.2} - 1)TE_{1.2}]$$

$$= 12 - [1.2975 + (3)(3.0740)]$$

$$= 12 - 10.5195$$

$$= 1.4805$$

and the actual leftover is

$LO_{1.2} = 12 - 10.5195 = 1.4805$.

So $LO_{1.2} = EXPLO_{1.2}$. It can be seen that $LO_{1.2} < TE_{1.2}$. The take every for sampling the second LTC PSU is $TE_{2.2} = \pi_2 ri$ and the random start is

$$RS_{2.2} = \left(\frac{TE_{1.2} - LO_{1.2}}{TE_{1.2}} \right) TE_{2.2}$$

$$= \left(\frac{3.0740 - 1.4805}{3.0740} \right) TE_{2.2}$$

$$= (.5184)\ TE_{2.2}, \text{ rounded to 4 decimal places}$$

e. Sample Counts

After sampling the 95+ eligible subfile, determine

$e_{i.2}$ = number of persons 95+ years of age and older sampled from the i^{th} LTC PSU in the 95+ eligible subfile.

B. Sample File

Following the sample selection, create a sample file for the 95+ supplement consisting of the entire compressed record (see Attachment B), plus the LTC PSU number and all sample identifiers and parameters listed in section X for each person selected for the 95+ supplement from the 95+ eligible subfile.

X. SAMPLE IDENTIFICATION CODES AND PARAMETERS

Several codes and parameter values are to be assigned to each person on the sample file. Some codes are to be applied in a systematic fashion. It is important that these codes are applied to sample persons in the prescribed order. Such ordering is specified below.

Assign the following codes and parameter values to the sample persons.

A. LTC Control Number

Each sample person is to be assigned a 14 digit LTC control number. The control number has the following form:

Digit 1:	screener or nonscreener code
Digit 2:	sample identification code 1 (based on status in 1982, also referred to as the 1984 sample component)
Digit 3:	sample identification code 2 (based on status in 1984, also referred to as the 1989 sample component)
Digit 4:	sample identification code 3 (based on status in 1989, also referred to as the 1994 sample component)
Digit 5:	sample identification code 4 (based on status in 1994 for longitudinal cohorts and identifies the aged-in and 95+ supplemental components for 1999)
Digits 6-8:	1970 CS PSU number (given in Attachment A)
Digits 9-12:	sequential numbers within 1970 CS PSUs
Digit 13:	healthy supplement code 1994 (identifies persons selected to receive the healthy supplement in 1994 if they are determined to be healthy)
Digit 14:	healthy supplement code 1999 (identifies persons selected to receive the healthy supplement in 1999 if they are

determined to be healthy)

The control number differs from the control number for the 1994 LTC survey by the addition and deletion of the following digits: the deletion of the zero-filled pad (digit 2), the addition of the sample identification code 4 (digit 5), and the addition of the healthy supplement code for 1999 (digit 14).

Below are instructions for determining each component of the control number.

1. Screener or Nonscreener Code (Digit 1)

   The first digit of a person's control number indicates whether the person receives a full or partial screener interview. The letter S indicates a full screener interview, and the letter N indicates a partial screener interview.

   Assign all persons selected from the 1999 universe subfile the letter S for digit 1.

2. Sample Identification Code 1 (Digit 2)

   The second digit of a person's control number will designate the sample group the person was from for the 1984 LTC survey. It is actually based on their status in 1982.

   Assign all persons selected from the 1999 universe subfile a 0 for digit 2.

3. Sample Identification Code 2 (Digit 3)

   The third digit of a person's control number will designate the sample group the person was from for the 1989 LTC survey. It is actually based on their status in 1984.

   Assign all persons selected from the 1999 universe subfile a 0 for digit 3.

4. Sample Identification Code 3 (Digit 4)

   The fourth digit of a person's control number will designate the sample group the person was from for the 1994 LTC survey. It is actually based on their status in 1989.

   Assign all persons selected from the 1999 universe subfile a 0 for digit 4.

5. Sample Identification Code 4 (Digit 5)

The fourth digit of a persons control number will designate the sample group the person is from for the 1999 LTC survey. For the longitudinal portion of the LTC sample it is actually based on a person's 1994 status. Persons selected by the instructions in this memorandum were not in the 1994 LTC survey.

Assign all persons selected for the aged-in cohort a 6 for this code. And assign all persons selected for the 95+ supplemental cohort a 9 for this code.

6. 1970 CS PSU Number (Digits 6-8)

Fill the sixth, seventh, and eighth digits of a person's control number with the 1970 CS PSU number of the LTC PSU the person was selected from. Note that the CS PSU numbers are given in Attachment A and are based on the 1970 design (the current survey design when LTC began). In LTC PSUs consisting of more than one CS PSU, use the CS PSU number given in Attachment C for the person's county, which is listed under his/her LTC PSU.

7. Sequential Numbers within 1970 CS PSUs (Digits 9-12)

Within each CS PSU, assign 4 digit numbers sequentially to each sample person in order of selection, starting with 0001 for each CS PSU.

8. Healthy Supplement Code -- 1994 (Digit 13)

Assign all persons selected for the 1994 healthy supplement the letter H for digit 13 of their control number. Assign all other persons the letter O for digit 13 of their control number.

9. Healthy Supplement Code -- 1999 (Digit 14)

Assign all persons selected for the 1999 healthy supplement the letter H for digit 14 of their control number. Assign all other persons the letter O for digit 14 of their control number.

Always write the control number with dashes after the sample identification code 4, after the CS PSU number, and before the healthy supplement code--1994. For example, S0006-105-0045-HH.

B. Reduction Group Code

Reduction group codes are assigned in case a sample needs to be reduced in the future. A sample may need to be reduced because of budgetary problems. The reduction group code should be assigned separately to the aged-in cohort and the 95+ supplement.

Assign a three digit code to each sample person selected for the aged-in cohort. Consider each aged-in cohort person as a string of persons in order of selection. Assign a reduction group code to each person, beginning with 003 for the first sample person, increasing by one until a person is assigned reduction group code 101, then continuing with 001 through 101 as often as necessary.

Assign a three digit code to each sample person selected for the 95+ supplement. Consider each person in the 95+ supplement as a string of persons in order of selection. Assign a reduction group code to each person, beginning with 121 for the first sample person, increasing by one until a person is assigned reduction group code 202, then continuing with 102 through 202 as often as necessary.

C. Stratum Code

Retain each person's stratum code, assigned in IV.A.1.

D. Quantities to be Used in Weighting the Sample

Assign the appropriate values of the quantities below to each sample person. The value of each varies by LTC PSU and/or stratum.

1. $TE_{i.1}$ for the aged-in cohort
   $TE_{i.2}$ for the 95+ supplement

2. π_i

3. F_{ij} = inflation factor for i^{th} LTC PSU and j^{th} stratum

   For the aged-in cohort,

   If $w_{ij1} > N_{ij1}$, then

   $$F_{ij} = \frac{w_{ij1}}{N_{ij1}},$$

Else F_{ij} = 10, for Group A PSUs
= 2, for Group B PSUs.

For the 95+ supplement,

If $w_{i.2} > E_{i.2}$, then

$$F_i = \frac{w_{i.2}}{E_{i.2}} ,$$

Else $F_i = 1$.

XI. OUTPUT

Certain printed output and tape files are needed upon completion of sample selection. Also, some tape files are needed prior to sample selection.

A. Universe Subfiles and the 95+ Eligible Subfile

Universe subfiles and the 95+ eligible subfile should be saved for possible use after sampling.

B. Sample Files

1. After selecting the aged-in cohort, create a sample file as specified in VI.B. Provide this file in the order that it was selected. For each sample person, this file should contain all items listed in Attachment B, in addition to the following:

   a. LTC PSU number
   b. LTC control number
   c. Reduction group code
   d. Stratum code
   e. $TE_{i.1}$ and π_i and F_{ij}

2. After selecting the 95+ supplemental sample, create a sample file as specified in IX.B. For each sample person, this file should contain all items listed in Attachment B, in addition to the following:

   a. LTC PSU number

b. LTC control number
c. Reduction group code
d. Stratum code
e. $TE_{i.2}$ and π_i and F_i

C. Sampling Data

1. Provide the following data to DSMD so that the sample selection can be verified:

   a. N_{ij1}

   b. $N_{i.2}$ $E_{i.2}$ across strata

   c. sampling intervals: si (aged-in) and ri (95+ supplement)

2. Provide the following data from the sample selection of the aged-in cohort from the first 1999 universe subfile. Data should appear in the order in which the sample was selected. Include in one file the following variables:

   a. π_i, $TE_{i.1}$, N_{ij1}, $N_{i.1}$, $N_{.j1}$, and $N_{..1}$ for each LTC PSU.

   b. RS_{ij1}, LO_{ij1}, $EXPLO_{ij1}$, α_j, and the first and last control numbers assigned, for each LTC PSU and stratum

   c. $\hat{n}_{ij1}$, n_{ij1}, $\hat{n}_{i.1}$, $n_{i.1}$, $\hat{n}_{.j1}$, $n_{.j1}$, $\hat{n}_{..1}$, $n_{..1}$

3. Provide a count of the number of persons designated for the healthy supplement (the number of persons who have an "H" in the 14th digit of their control number). This should be around 1,545.

4. Provide the following data from the sample selection of the 95+ supplement from the 95+ eligible subfile. Data should appear in the order in which the sample was selected. Include in one file the following variables:

   a. π_i, $TE_{i.2}$, $E_{i.2}$, and $E_{..2}$ for each LTC PSU.

   b. $RS_{i.2}$, $LO_{i.2}$, $EXPLO_{i.2}$, α_j, and the first and last control numbers