Appendix A

Sample Design, Weighting and Estimation Procedures, and Computation of Sampling Errors

Adam Chu, Carla E. Maffeo, Annie Lo, David Morganstein, Karen J. Bandeen-Roche, Judith D. Kasper, Dwight B. Brock

Sample Design

The sampling frame for the Women's Health and Aging Study (WHAS) was restricted to Medicare beneficiaries residing in 12 Zip Code areas in and adjoining the city of Baltimore, Maryland. Four nonoverlapping probability samples, spaced approximately 6 months apart, were selected over a 2-year period. These four samples are referred to as "replicates" and are denoted by number in this appendix (i.e., replicate 1, replicate 2, etc.). The sample design for each replicate of the WHAS can be described as a stratified random sample of female Medicare beneficiaries with primary strata defined by the following three age groups: (1) 65 to 74 years; (2) 75 to 84 years; and (3) 85 years and older.

The sampling frame of Medicare beneficiaries in the target population was constructed from current Health Care Financing Administration (HCFA) files. To facilitate sample selection and updating in future replicates of the study, the original sampling frame was randomly divided into four mutually exclusive subsets that were balanced with respect to age group and Zip Code. As detailed below, the first of these subsets was then used to select the sample of beneficiaries for replicate 1. The remaining subsets were set aside for use in the subsequent replicates of the study and were updated for deaths, moves, and new enrollees using the most recent information available at the time of sample selection. Beneficiaries who were previously included in the Medicare Current Beneficiary Survey (MCBS) or the Senior Health Watch Study were deleted from the WHAS frame before sample selection, since these studies were also being conducted in Baltimore and there was a concern about respondent burden and potential nonresponse. It should be noted that since the samples for the other two studies were selected randomly, no sampling bias is introduced by excluding these persons from WHAS. Table A.1 summarizes the sample sizes for each of the four WHAS replicates by age group. The sample sizes shown in Table A.1 are for the "initial" samples of beneficiaries who were screened for eligibility. The corresponding numbers of beneficiaries who were eligible for and completed the full baseline interview and nurse's examination are shown in Table A.4. In the remainder of this appendix, women who completed the full baseline interview and nurse's examination will be denoted as completing the examination. Additional information about the procedures used to select each of the four replicate samples is given below.

The sampling frame for the first replicate of the WHAS included 32,538 female Medicare beneficiaries who resided in the 12 Zip Code areas specified for the study and who were age 65 years or older as of September 1, 1992. Of the 32,538 eligible beneficiaries, 2,274 were MCBS or Senior Health Watch participants and were deleted from the sampling frame. The remaining 30,264 eligible beneficiaries were then stratified by Zip Code within each of three age groups (65 - 74, 75 - 84, and 85+) and randomly (and systematically) assigned to one of four mutually exclusive subsets of approximately equal size. The 7,566 beneficiaries who were assigned to subset 1 were retained for replicate 1. Subsets 2 to 4 were set aside for future use in subsequent replicates of the study.

Table A.1.Initial Sample Sizes for the WHAS, by Replicate and Age Group

Replicate	Age group (at time of sampling)			Total
	65 to 74 years	75 to 84 years	85 years and older	
1. September 1, 1992	600	425	340	1,365
2. March 1, 1993	600	425	340	1,365
3. September 1, 1993	600	476	571	1,647
4. March 1, 1994	752	481	911	2,144
Total	2,552	1,807	2,162	6,521

Within the subset designated for the first replicate, 600 beneficiaries were selected from the 65 to 74 years age group, 425 from the 75 to 84 years age group, and 340 from the age 85 years and older group, for a total sample size of 1,365 beneficiaries. The sample sizes established for each age group were designed to yield a total sample of about 250 study participants who met specific physical disability and cognitive function criteria (Kasper and Rohde, 1992; Shapiro, 1991). Within each age group, the samples were selected at rates that varied by Zip Code to take account of the previous selection of MCBS and Senior Health Watch participants. However, the overall probabilities of selection were independent of Zip Code for each age group; that is, the sample was self-weighting within each of the three age groups but not across the three age groups.

The sampling frame for the second replicate consisted of the 31,938 eligible female beneficiaries residing in the 12 target Zip Code areas who were age 65 years or older as of March 1, 1993 (i.e., an updated frame was used for the second replicate). The selection of the sample for this replicate proceeded as follows. First, the subset of 7,566 previously enrolled beneficiaries who had been designated for replicate 2 (i.e., subset 2 created during the selection of the sample for replicate 1) was updated for deaths and moves, which reduced the number of beneficiaries to 6,897. Next, a frame of new beneficiaries was constructed by matching the most recent HCFA file against the earlier HCFA file used to select the sample for replicate 1. This matching process identified 4,300 new beneficiaries who were then randomly assigned to one of four subsets of equal size. Subset 2, consisting of 1,075 new beneficiaries, was retained for further subsampling.

However, it was later discovered that the 4,300 "new" beneficiaries identified above also included MCBS and Senior Health Watch participants, who should have been excluded from this set. (This problem was corrected in subsequent replicates of the study.) Of the 1,075 "new" beneficiaries assigned to subset 2, only 552 were actually new enrollees. The rest were MCBS and Senior Health Watch participants. Since the inadvertent inclusion of MCBS and Senior Health Watch participants was not discovered at the time of sample selection, the file from which the sample for the second replicate was selected consisted of 6,897 + 1,075 = 7,972 beneficiaries. From this subset, 600 beneficiaries were selected from the age 65 to 74 years group, 425 from the age 75 to 84 years group, and 340 from the age 85 years and older group (age was determined as of March 1, 1993). Generally, any MCBS or Senior Health Watch participants who were sampled as part of the "new enrollee" sample were retained for the study.

Because the target sampling rates for the second replicate were derived under the erroneous assumption that MCBS and Senior Health Watch participants were excluded from the set of 4,300 newly enrolled beneficiaries, the actual overall sampling rates for the previously enrolled beneficiaries in replicate 2 varied slightly by Zip Code. However, this variation in sampling rates was not expected to seriously inflate the sampling errors of estimates from the survey.

The sampling frame for the third replicate consisted of the 31,068 eligible female beneficiaries residing in the 12 target Zip Code areas who were age 65 years or older as of September 1, 1993. Before the sample for this replicate was selected, the subset of 7,566 previously enrolled beneficiaries designated for replicate 3 (i.e., subset 3 created during the selection of the sample for replicate 1) was updated for deaths and moves. This reduced the number of previously enrolled beneficiaries in this subset to 6,497. Next, a frame of new beneficiaries was constructed in two parts. The subset of new beneficiaries (subset 3) identified during the creation of the frame for the second replicate was updated for deaths and moves, which resulted in a subset of 497 new beneficiaries who were identified in the second data collection period and who survived to the current period. An additional 1,126 new beneficiaries were identified by matching the most recent HCFA file against the earlier HCFA files used to select the samples for replicates 1 and 2. The 1,126 newly identified beneficiaries were then randomly assigned to one of four subsets of approximately equal size. Subset 3, consisting of 281 of these new beneficiaries, was retained for further subsampling.

The file from which the sample for the third replicate was selected consisted of 6,497 + 497 + 281 = 7,275 beneficiaries. From this subset, 600 beneficiaries were selected from the age 65 to 74 years group, 476 from the age 75 to 84 years group, and 571 from the age 85 and older group (age was determined as of September 1, 1993), for a total sample of 1,647 beneficiaries. The somewhat larger sample sizes specified for replicate 3 were intended to compensate for the lower-than-expected study yields in the previous two replicates. The within-Zip Code sampling rates used to select the previously enrolled beneficiaries varied by Zip Code to compensate for the exclusion of MCBS and Senior Health Watch participants. However, the resulting overall probabilities of selection were independent of Zip Code within age group for both previous and new beneficiaries.

The sampling frame for the fourth and final replicate included 31,488 eligible female beneficiaries residing in the 12 target Zip Code areas who were age 65 years or older as of March 1, 1994. Before the sample for this replicate was selected, the subset of 7,564 previously enrolled beneficiaries designated for replicate 4 (i.e., subset 4 created during the selection of the sample for replicate 1) was updated for deaths and moves, which reduced the number of previously enrolled beneficiaries in this subset to 6,255. Next, a frame of new beneficiaries was constructed in three parts. The subset of new beneficiaries (subset 4) identified during the creation of the frame for the second replicate was updated for deaths and moves, which resulted in a subset of 488 new beneficiaries who were identified in the second data collection period and who survived to the current period. The subset of new beneficiaries (subset 4) identified during the creation of the frame for the third replicate was also updated for deaths and moves, which resulted in a subset of 255 new beneficiaries who were identified in the third data collection period and who survived to the current period. An additional 1,646 new beneficiaries were identified by matching the most recent HCFA file against the earlier HCFA files used to select the samples for replicates 1, 2, and 3. The 1,646 newly identified beneficiaries were then randomly assigned to one of four subsets of approximately equal size. Subset 4, consisting of 411 of these new beneficiaries, was retained for further subsampling.

The file from which the sample for the fourth replicate was selected therefore consisted of 6,255 + 488 + 255 + 411 = 7,409 beneficiaries. From this subset, 752 beneficiaries were selected from the age 65 to 74 years group, 481 from the age 75 to 84 years group, and 911 from the age 85 years and older group (age was determined as of March 1, 1994), for a total sample of 2,144 beneficiaries. The sample sizes specified for the fourth replicate reflected additional adjustments designed to more closely achieve the study's overall sample size goals within age groups. The within-Zip Code sampling rates used to select the previously enrolled beneficiaries varied by Zip Code to compensate for the exclusion of MCBS and Senior Health Watch participants. As a result of the larger sample size requirements for the oldest age group, all available beneficiaries designated for the fourth replicate in some Zip Codes were included in the sample. The resulting overall probabilities of selection therefore varied slightly by Zip Code for the oldest age group.

Within the two younger age groups, the overall probabilities of selection were independent of Zip Code.

Weighting and Estimation Procedures

The estimates of means and proportions presented in this monograph were calculated using weights that inflate the respondent data to population levels. Such weights are needed to properly reflect sample design features such as stratification and variable probabilities of selection, and also to compensate for differential nonresponse rates (e.g., see Skinner et al., 1989). As described below, two sets of weights were developed for the analysis of the WHAS data, including one set for the initial (screening) sample and another for the final study sample (women who completed both the baseline interview and the followup nurse's examination). The procedures used to construct the weights for the WHAS samples are described in the following sections.

Weighting the Initial (Screener) Sample

The first step in the weighting process was to assign base weights equal to the reciprocals of the overall probabilities of selection to each beneficiary included in the initial sample. The sum of the base weights (when summed over all beneficiaries in the screening sample) provides an unbiased estimate of the number of beneficiaries in the HCFA frame at the time the sample was selected. The average base weights assigned to the sampled beneficiaries and the corresponding weighted sample counts are summarized in Table A.2 by age group and replicate.

Table A.2. Base Weights Assigned to Sampled Beneficiaries, by Replicate and Age Group

Replicate	Age group (at time of sampling)			Weighted sample count
	65 to 74 years	75 to 84 years	85 years and older	
1. September 1, 1992	26.78	27.84	13.63	32,534
2. March 1, 1993	27.23*	26.42*	12.71*	31,888
3. September 1, 1993	25.74	23.80	7.53	31,072
4. March 1, 1994	21.15	23.41	4.75*	31,492

*Owing to varying probabilities of selection (see section on sampling), entry corresponds to average weight of sampled persons in the given age group and replicate.

To compensate for losses owing to screener nonresponse, the base weights were adjusted within broad classes defined by age group, race, and geography (Zip Code). Collapsing across Zip Codes was often necessary to ensure a minimum sample size of about 15 to 20 beneficiaries in each final weighting class. To calculate the required nonresponse adjustments, the sampled beneficiaries were assigned to one of the four screener response-status groups defined in Table A.3. Note that the screener nonrespondents were classified into one of two groups depending on their presumed eligibility for the screener. The type 1 nonrespondents included nonrespondents who were known not to have moved, been institutionalized, or died, while the type 2 nonrespondents included nonrespondents who may have moved, been institutionalized, or died.

Conceptually, the nonresponse adjustments were made in two stages. At the first stage of adjustment, the total weight of the type 2 nonrespondents in weighting class h was distributed in proportion to the remaining groups in the sample; that is, an initial adjusted weight for the i th screener respondent in class h was calculated as:

$$w_{hi}^{(1)} = w_{hi}^{base}\left(\frac{S_1 + S_2 + S_3 + S_4}{S_1 + S_3 + S_4}\right), \quad (1)$$

where w_{hi}^{base} is the base weight for the i th screener respondent in class h, and S_k is the sum of the base weights, summed over the n_{hk} sampled beneficiaries in response-status group k (k = 1, 2, 3, 4), where the four response-status groups are defined in Table A.3. In effect, a proportion of the type 2 nonrespondents was treated as eligible for the screener survey (i.e., have not moved, become institutionalized, or died), and the complementary proportion was considered to be out of scope (i.e., have moved, become institutionalized, or died).

At the second stage of adjustment, the previously adjusted weights of the screener respondents were further inflated to compensate for the type 1 respondents; that is, the final screener nonresponse-adjusted weight for the i th screener respondent in class h (whether or not the respondent qualified for the full baseline interview and nurse's examination) was calculated as

$$w_{hi}^{NR} = w_{hi}^{(1)}\left(\frac{S_1^* + S_2^*}{S_1^*}\right), \quad (2)$$

where S_k^* is the sum of the $w_{hi}^{(1)}$'s, summed over the n_{hk} sampled beneficiaries in response- status group k (k = 1, 2).

Note that the w_{hi}^{NR}'s defined by formula (2) are the appropriate weights for analyzing the screener survey data for any particular replicate.

Table A.3. Distribution of Sampled Persons by
Screener Response Status and Replicate

Screener response status group	WHAS replicate				Total
	1	2	3	4	
1. Respondents: persons for whom a completed screener was obtained, whether or not the person qualified for the full baseline interview and nurse's exam	883	903	1,068	1,283	4,137
2. Type 1 Nonrespondents: persons for whom a completed screener was not obtained, but who have not died, been institutionalized, or moved.	169	210	257	380	1,016
3. Type 2 nonrespondents: persons for whom a completed screener was not obtained, but who may have died, been institutionalized, or moved.	33	28	38	64	163
4. Out of scope (ineligible for the screener): persons who died, were institutionalized, or moved out of the survey area.	280	224	284	417	1,205
Total unweighted count	1,365	1,365	1,647	2,144	6,521
Total weighted count of respondents (response status group 1) using final screener weights	26,506	27,230	27,016	27,309	---

Weighting the Examination Sample

Ordinarily, the weight for a person for whom an examination was conducted is equal to the nonresponse-adjusted weight, w_{hi}^{NR} . However, for various reasons, not all of those who qualified for the full baseline interview and nurse's examination completed the full assessment. To compensate for the examination nonrespondents, an additional adjustment was made within classes defined by age group, number of domains of disability (2, 3, or 4), and Mini-Mental State Examination score (less than 25 or 25 or higher).

Specifically, let n_{g1} denote the number of persons in adjustment class g for whom examination data were obtained (examination "respondents"), and let n_{g2} denote the corresponding number of persons who qualified for the examination, but for whom examination data were not obtained (examination "nonrespondents"). The final examination weight for the i th respondent in adjustment class g was computed as

$$w_{gi}^{exam} = w_{gi}^{NR}\left(\frac{S_1^{**} + S_2^{**}}{S_1^{**}}\right), \quad (3)$$

where S_k^{**} is the sum of the w_{gi}^{NR}'s, summed over the n_{gk} sampled beneficiaries in examination response-status group k (k = 1, 2). Table A.4 summarizes the numbers of examination respondents by replicate and age group, along with the corresponding weighted counts of respondents using the final examination weights.

Table A.4. Distribution of Examination Respondents by Replicate and Age Group*

Replicate	Unweighted count of baseline interview and examination respondents			Weighted sample count
	65 to 74 years	75 to 84 years	85 years and older	
1. September 1, 1992	89	79	44	8,611
2. March 1, 1993	103	77	43	8,766
3. September 1, 1993	77	77	89	7,796
4. March 1, 1994	119	78	127	9,554
Total	388	311	303	--

*Age group is based on the survey-reported age.

Estimates for All WHAS Replicates Combined

The sample-based estimates presented in this monograph were obtained by combining the weighted results from all WHAS replicates. The combined estimate therefore represents a weighted average of the corresponding estimates for each of the four replicates. Specifically, let $\bar{x}_t$ denote the estimated mean value of a survey item, X, for WHAS replicate t ; that is,

$$\bar{x}_t = \frac{\sum_{i=1}^{n_t} w_{ti}^{exam} x_{ti}}{\sum_{i=1}^{n_t} w_{ti}^{exam}}, \quad (4)$$

where x_{ti} is the observed value of X for respondent i in WHAS replicate t, w_{ti}^{exam} is the corresponding sampling weight for respondent i in WHAS replicate t, and n_t is the sample size (number of respondents) for WHAS replicate t. The corresponding estimate for all replicates combined, $\bar{x}_{comb}$, was then computed as

$$\bar{x}_t = \frac{\sum_{t=1}^{4}\sum_{i=1}^{n_t} w_{ti}^{exam} x_{ti}}{\sum_{t=1}^{4}\sum_{i=1}^{n_t} w_{ti}^{exam}}, \quad (5)$$

It should be noted that $\bar{x}_{comb}$ provides an unbiased estimate of the average population mean

$$\mu = \frac{\sum_{t=1}^{4} N_t \mu_t}{\sum_{t=1}^{4} N_t}, \quad (6)$$

where μ_t is the mean value of X for the eligible population of beneficiaries at time t (WHAS replicate t), and N_t is the corresponding size of the eligible population at time t. Each $\bar{x}_t$ estimates μ_t. In the 2-year period during which the WHAS was conducted, the N_t's did not vary importantly from replicate to replicate; thus, for all practical purposes, the weighting factor,

$$N_t \Big/ \sum_{t=1}^{4} N_t$$

was approximately 1/4 for all t.

Computation of Sampling Errors

Because the sample design for the WHAS was a stratified probability sample, variance estimation based on the assumption of simple random sampling is not appropriate (for example, see Skinner et al., 1989). To properly reflect design features used in the WHAS such as stratification and systematic sampling, the sampling errors (or variances) of the survey-based estimates were calculated by a pseudo-replication method known as jackknife replication. Under jackknife replication, a specified number of systematic subsamples were generated from the full sample, and these in turn were used to define a series of jackknife replicates[1] by dropping one subsample at a time from the full sample. Each jackknife replicate was then reweighted using the weighting procedures developed for the full sample, and the resulting replicate-specific weights were attached to each data record to facilitate variance estimation. The advantage of the jackknife replication method is that it provides a relatively simple way of calculating the sampling errors of estimates from a complex sample design (for example, see McCarthy, 1966; Wolter, 1985).

Each jackknife replicate was formed as follows. The data from the selected women were first sorted by time period (i.e., by WHAS replicate). Within each time period they were sorted by age group and then by Zip Code within age group; that is, they were arranged in their sample selection order. A jackknife replicate was defined by leaving out every 31st sampled woman and increasing the weights of those retained by 31/30 so the weights add to the correct total. The rth jackknife replicate consisted of everyone *except* the rth, r + 31th, 2r + 31th, etc., women.

To illustrate how the sampling errors were computed, let $\bar{x}$ denote a weighted mean or proportion based on the full WHAS sample. Further, let $\bar{x}^{(r)}$ denote the corresponding estimate based on jackknife replicate r. The estimated variance of $\bar{x}$ was then computed from the formula

$$\text{var}(\bar{x}) = \left(\frac{R-1}{R}\right)\sum\left(\bar{x}^{(r)} - \bar{x}\right)^2 , \quad (7)$$

where the summation extends over the R jackknife replicates defined for variance estimation. In practice, R is usually designed to be between 30 and 50; for the WHAS, R was set to 31. Note that the square root of $\text{var}(\bar{x})$ is the standard error of $\bar{x}$.

WHAS Variance Estimation

Although the jackknife replication technique makes the estimation of the sampling variance of any statistic straightforward, the estimation process is computationally intensive. Standard statistical software does not provide a method for performing these computations; accordingly, it is necessary to use special-purpose programs in addition to programs that perform the analyses or tabulations. To reduce the work required to calculate sampling errors and to reduce the size of the publishing task for each estimate, an approximation of the standard error of an estimated

[1] The term "jackknife replicate" should not be confused with the four "replicates" defined for WHAS. As described in this appendix, a jackknife replicate is simply a specially constructed subsample of the full WHAS sample.

population proportion or mean is frequently used. With the approximate method, analysts can use simple formulas to obtain approximate standard errors from the estimates themselves, while still accounting for the effects of a complex sample design.

To approximate the standard error for an estimated population proportion or mean, design effects based on the coefficient of variation of the sampling weights were calculated. A design effect expresses the efficiency of the design compared to simple random sampling and is defined to be the ratio of the variance of the estimate obtained from the WHAS sample to the variance of the estimate obtained from a simple random sample of the same sample size (Kish, 1965). Under simple random sampling, the variance of an estimate of a proportion is $p(1-p)/n$, where p is the proportion of the sample having the characteristic, and n is the sample size used in calculating the proportion. Thus, the standard error for an estimated proportion is approximately given by

$$SE(\hat{p}) = \sqrt{DEFF \frac{\hat{p}(1-\hat{p})}{n}}, \quad (8)$$

where SE is the standard error, $\hat{p}$ is the estimated population proportion, and $DEFF$ is the design effect.

Since the WHAS sample design was an unclustered, stratified probability sample design using sampling rates that varied by stratum and replicate (time), the design effect can be computed easily from the coefficient of variation of the sampling weights. For any subgroup of the sample, the $DEFF$ of the WHAS design is $DEFF = 1 + (cv_w)^2$ where cv_w is the coefficient of variation of the weights (Kish, 1992). Table A.5 shows the design effects due to variable weights, separately for the WHAS screener and the examination samples by age group.

Table A.5. Design Effects by Sample and Age Group

Age group (survey-reported)	Screener	Examination
65 to 74 years	1.013	1.019
75 to 84 years	1.015	1.024
85 years and older	1.313	1.302
Total	1.123	1.159

To assess the adequacy of the standard error approximation, unbiased estimates of variances using the jackknife replication technique were computed for a large number of statistics. The statistical software procedure WESVAR (Westat, 1989) was used to compute the standard errors for estimates related to functioning domains, measured walks, functional reach, repeated chair stands, and walking aids. The approximate standard errors given by formula (8) were found to be comparable to the standard errors generated by the jackknife replication technique. Thus, the approximation provides an easy and efficient way for analysts to assess the sampling precision of the survey-based estimates presented in the monograph.

An example for computing the standard error for a percentage is illustrated using data on walking aids. For the examination sample, an estimated 39.9 percent of the 311 Medicare beneficiaries age 75 to 84 years reported they used a cane when they walked (i.e., $\hat{p} = 0.399$, where $\hat{p}$ is the weighted estimate). Using the design effect of 1.024 from Table A.5, the standard error can be computed as follows:

$$SE(\hat{p}) = \sqrt{\frac{(0.399)(0.601)}{311}(1.024)} = 0.0281 \text{ or } (2.81\%).$$

A 95 percent confidence interval for the percentage of persons age 75 to 84 years who use a cane can be constructed as

$$\hat{p} \pm 1.96SE(\hat{p})$$

Substituting 39.9 percent for $\hat{p}$ and 2.81 percent for $SE(\hat{p})$, a 95 percent confidence interval for the percentage P of Medicare beneficiaries age 75 to 84 years who use a cane when they walk is 39.9 ± 1.96(2.81), or $34.4 < P < 45.4$.

The 95 percent confidence interval can be interpreted as follows: under repeated sampling with the same sample design, approximately 95 percent of intervals constructed as above will contain the population value of P.

This procedure also can be used to compute standard errors and confidence intervals for means. The standard error for an estimated mean, $\bar{x}$, can be approximated using the following formula:

$$SE(\bar{x}) = \sqrt{DEFF\frac{s^2}{n}}, \quad (9)$$

where $\bar{x}$ is the estimated (weighted) mean of the variable x, $DEFF$ is the design effect shown in Table A.5, and s^2 is an estimate of the population variance of x. For example, s^2 can be estimated using the formula

$$\frac{\sum w_i\left[x_i-\left(\frac{\sum w_i x_i}{\sum w_i}\right)\right]^2}{\sum w_i} \quad (10)$$

References

Kasper J, Rohde C. (1992). *Sample Design for the Women's Health and Aging Study*, The Johns Hopkins University, unpublished memorandum, 5/8/92.

Kish L. (1965). *Survey Sampling*. New York: Wiley & Sons, p. 88.

Kish L. (1992). Weighting for unequal P_i. *J Official Stat* 8:183-200.

McCarthy P. (1966). Replication: An approach to the analysis of data from complex surveys. *Vital Health Stat*. Series 2, No. 14, U.S. Department of Health, Education, and Welfare. Washington, DC.

Shapiro S. (1991). *Sample Design: Women's Health and Aging Study*, The Johns Hopkins University, unpublished procedural memorandum, 2/23/91.

Skinner C, Holt D, Smith T, eds. (1989). *Analysis of Complex Surveys*. New York: Wiley & Sons.

Westat. (1989). WESVAR documentation, Rockville, MD.

Wolter K. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.