
Program Memorandum Carriers

Department of Health & Human
Services (DHHS)
Centers for Medicare &
Medicaid Services (CMS)

Transmittal B-03-022

Date: MARCH 21, 2003

This Program Memorandum re-issues Program Memorandum B-02-007, Change Request 1363 dated February 7, 2002. The only change is the discard date; all other material remains the same.

CHANGE REQUEST 1363

SUBJECT: Use of Statistical Sampling for Overpayment Estimation When Performing Administrative Reviews of Part B Claims

This Program Memorandum (PM) provides clarified guidance and direction for Medicare carriers to use when conducting statistical sampling for overpayment estimation. The attached replaces the prior Sampling Guidelines Appendix for reviews conducted after issuance of this PM. For reviews conducted prior to this issuance, the attached are a clarification to aid interpretation of the earlier instructions, particularly where specific numbers are suggested.

This PM obsoletes the Medicare Carriers Manual (MCM) Sampling Guidelines Appendix that has been in effect since December 1975 (MCM Part 3, Chapter VII).

The *effective date* for this PM is January 8, 2001.

The *implementation date* for this PM is as soon as possible, but no later than February 9, 2001.

This PM should be implemented within your current operating budget.

This PM may be discarded after March 20, 2004.

If you have any questions, contact Betsy Horn at (410) 786-0973.

Attachment

USE OF STATISTICAL SAMPLING FOR OVERPAYMENT ESTIMATION WHEN PERFORMING ADMINISTRATIVE REVIEWS OF PART B CLAIMS

I. Introduction.

A. General.--The purpose of these instructions is to provide Medicare carriers (hereinafter you or your) with the guidance necessary to use statistical sampling to calculate and project overpayments identified following administrative review of Part B claims. These instructions are provided to ensure that a statistically valid sample is drawn and that statistically valid methods are used to project an overpayment where the results of the review indicate that overpayments have been made. These guidelines are for administrative reviews performed by you. Reviews using statistical sampling conducted by you to assist with the identification, case development and/or investigation of suspected fraud or other unlawful activities may use procedures that differ from those prescribed herein.

These instructions are provided so that you follow a sufficient administrative process when you are conducting statistical sampling to project overpayments. Failure by you to follow one or more of the requirements contained herein does not necessarily affect the validity of the statistical sample. An appeal challenging the validity of the sampling methodology must be predicated on the actual statistical validity of the sample as drawn and conducted. Your failure to follow one or more requirements may result in review by CMS of your performance, but should not be construed as necessarily affecting the validity of the statistical sample.

B. Use of Statistical Sampling.--Statistical sampling is used to calculate and project the amount of overpayments made on Part B claims. CMS Ruling 86-1 (CMSR 86-1) explains CMS's authority to use statistical sampling to estimate overpayments made to physicians and suppliers. The ruling recognizes that statistical sampling conserves the resources of the Medicare program when reviews are performed on a large universe of claims. The ruling states that in most cases it would not be administratively feasible, given the volume of records involved and the cost of retrieving and reviewing all the beneficiary records, for you to examine all individual claims for the period in question.

C. Statistical Sampling Steps.--The major steps in conducting statistical sampling are: (1) Selecting the physician or supplier; (2) Selecting the period to be reviewed; (3) Defining the universe, the sampling unit, and the sampling frame; (4) Designing the sampling plan and selecting the sample; (5) Reviewing each of the claims (or portions thereof), and determining if there was an overpayment, or, for administrative reviews, an underpayment; and, as applicable, (6) Estimating the overpayment. Where an overpayment has been determined to exist, follow applicable instructions for notification and collection of the overpayment (see §VII).

D. When Statistical Sampling is Appropriate.--You may use statistical sampling to project overpayments to physicians and suppliers when erroneous billing or reimbursement, or overutilization is suspected, and when a case-by-case review is not administratively feasible or practical.

Your use of statistical sampling to determine overpayments may be used in conjunction with other corrective actions. Reviews that involve the use of statistical sampling may be utilized when there is a "major level of concern" regarding the physician or supplier's billing, reimbursement, and/or utilization (see Program Memorandum AB-00-72, dated 8/7/2000, Change Request 1285 - Progressive Corrective Action (PCA)).

Factors also to be considered for determining when to undertake statistical sampling include, but are not limited to, the number of claims in the universe and the dollar values associated with those claims; your available resources; and the cost effectiveness of the expected sampling results.

E. Consultation With a Statistical Expert.--Initially, all statistical sampling procedures you use must be reviewed by a statistician or other person with equivalent expertise in statistical sampling and estimation methods. This is done to ensure that a statistically valid sample is drawn and that you use statistically valid methods. You must obtain from the statistical expert a written approval of the methodology for the type of statistical sampling to be performed. If this sampling methodology is applied routinely and repeatedly, the original written approval is adequate for conducting subsequent reviews utilizing the same methodology. You must have the statistical expert review the results of the sample prior to releasing the demand letter. If questions or issues arise during the on-going review, you must also involve the statistical expert.

At a minimum, the statistical expert you use (either on-staff or consultant) should possess a master's degree in statistics or have equivalent experience. See Appendix A for a list, not exhaustive, of texts that represent the minimum level of understanding that the statistical expert should have. If you do not have staff with sufficient statistical experience as outlined here, obtain such expert assistance prior to conducting statistical sampling.

F. Use of Other Sampling Methodologies.--Nothing in these instructions precludes the Health Care Financing Administration (CMS) or you from relying on statistically valid audit sampling methodologies employed by other audit organizations, including but not limited to the Office of Inspector General, the General Accounting Office, and other authoritative sources. Where it is foreseen that the results of your review may be referred to a law enforcement or other agency for litigation and/or other enforcement actions, discuss specific litigation and/or other requirements as they relate to statistical sampling with your statistical expert prior to undertaking the review. In addition, discuss sampling requirements with law enforcement or other authorities before initiating the review (to ensure that your review will meet their requirements, and that such work will be funded accordingly).

II. Probability Sampling.--

Regardless of the method of sample selection you use, the procedure must result in a probability sample. For a procedure to be classified as probability sampling the following features must apply:

- It must be possible, in principle, to enumerate a set of distinct samples that the procedure is capable of selecting if applied to the target universe. Although only one sample will be selected, each distinct sample of the set has a known probability of selection. It is not necessary to actually carry out the enumeration or calculate the probabilities, especially if the number of possible distinct samples is large - possibly billions. It is merely meant that one could, in theory, write down the samples, the sampling units contained therein, and the probabilities if one had unlimited time.
- Each sampling unit in each distinct sample has a known probability of selection. One of the possible samples is selected by a random process according to which sampling unit receives its appropriate chance of selection. The selection probabilities do not have to be equal but they should all be greater than zero. In fact, some designs bring gains in efficiency by not assigning equal probabilities to all of the distinct sampling units.

For a procedure that satisfies these bulleted properties it is possible to develop a mathematical theory for various methods of estimation based on probability sampling and to study the features of the estimation method (i.e., bias, precision, cost) although the details of the theory may be complex. If a particular probability sample design is properly executed, i.e., defining the universe, the frame, the sampling units, using proper randomization, accurately measuring the variables of interest, and using the correct formulas for estimation, then assertions that the sample and its resulting estimates are "not statistically valid" cannot legitimately be made. In other words, a probability sample and its results are always "valid." Because of differences in the choice of a design, the level of available resources, and the method of estimation, however, some procedures lead to higher precision (smaller confidence intervals) than other methods. A feature of probability sampling is that the level of uncertainty can be incorporated into the estimate of overpayment as is discussed below.

III. Selection of Period to be Reviewed and Composition of Universe.--

A. Selection of Period for Review.--Following your selection of the physician or supplier, determines the period of time to be reviewed. That is, determine the number of days, weeks, months, or years, for which sampling units will be reviewed. You will select your universe from this period. The period of review is determined by considering several factors, including (but not limited to):

- How long the pattern of erroneous billing or overutilization is believed to have existed;
- The volume of claims that are involved;
- The length of time that a national average decision or regional or local coverage policy has been in effect (i.e., should the physician or supplier have succeeded in adjusting their billing/utilization practices by now);
- The extent of prepayment review already conducted or currently being conducted;
- The dollar value of the claims that are involved relative to the cost effectiveness of the sample; and/or
- The applicable time periods for reopening claims (see the Medicare Carriers Manual, §§12100ff., for Reopening Standards.)

B. Defining the Universe, the Sampling Unit, and the Sampling Frame.--Your universe and sampling frame will usually be all relevant claims for the period under review. The discussion which follows assumes that the unit of the universe is the claim, although situations may arise in which it is necessary to review all claims for a beneficiary, or all claims on all patients treated on the same day. The unit of the universe may be the patient, a treatment “day”, or any other unit appropriate for the issue under review.

1. Composition of the Universe.--The universe will be all fully and partially paid claims submitted by the physician or supplier for the defined period of review for the sampling units to be reviewed. For example, if you are reviewing Physician X for the period January 1, 2000 – March 31, 2000, and you have selected for review laboratory and other diagnostic tests, your universe would include all fully and partially paid claims for laboratory and diagnostic tests billed by that physician for the selected time period. For some reviews the period of review may be best defined in terms of the date of service because changes in coverage policy may have occurred.

2. The Sampling Unit.--Sampling units are the elements that are selected according to the chosen method of statistical sampling. They may be individual lines within claims, individual claims, or clusters of claims (e.g., a beneficiary). For example, possible sampling units may include specific beneficiaries seen by a physician during the time period under review; or, claims for a specific item or service. In certain circumstances, e.g., multi-stage sample designs, other types of clusters of payments may be used. In principle, any type of sampling unit is permissible as long as the total aggregate of such units covers the population of potential mispaid amounts.

3. The Sampling Frame.--The sampling frame is the “listing” of all the possible sampling units from which the sample is selected. The frame may be, for example, a list of all beneficiaries receiving items from a selected supplier, a list of all claims for which fully or partially favorable determinations have been issued, or a list of all the line items for specific items or services for which fully or partially favorable determinations have been issued.

The ideal frame is a list that covers the target universe completely. In some cases the frame must be constructed by combining lists from several sources and duplication of sampling units may result. Although duplicate listings can be handled in various ways that do not invalidate the sample, it is recommended that you eliminate the duplicates before selecting the sample.

IV. Sample Selection.

A. Sample Design.--Identify your sample design. The most common designs used are simple random sampling, systematic sampling, stratified sampling, and cluster sampling, or a combination of these.

1. Simple Random Sampling.--Involves using a random selection method to draw a fixed number of sampling units from the frame without replacement, i.e., not allowing the same sampling unit to be selected more than once. The random selection method must ensure that, given the desired sample size, each distinguishable set of sampling units has the same probability of selection as any other set - thus the method is a case of "equal probability sampling." An example of simple random sampling is that of shuffling a deck of playing cards and dealing out a certain number of cards (although for such a design to qualify as probability sampling a randomization method that is more precise than hand shuffling and dealing would be required.)

2. Systematic Sampling.--Requires that the frame of sampling units be numbered, in order, starting with the number one (1) and ending with a number equal to the size of the frame. Using a random start, the first sampling unit is selected according to that random number, and the remaining sampling units that comprise the sample are selected using a fixed interval thereafter. For example, if a systematic sample with size one-tenth of the frame size is desired, select a random number between one and ten, say that it is "6", and then select every tenth unit thereafter, i.e., "16, 26, 36," until the maximum unit number in the frame has been exceeded.

3. Stratified Sampling.--Involves classifying the sampling units in the frame into non-overlapping groups, or strata. One useful stratification results in a sampling unit from one stratum more likely being similar in overpayment amount to others in its stratum than to sampling units in other strata. Although the amount of an overpayment cannot be known prior to review, it may be possible to stratify on an observable variable that is correlated with the overpayment amount of the sampling unit. Given a sample in which the total frame is covered by non-overlapping strata, if independent probability samples are selected from each of the strata, the design is called stratified sampling. The independent random samples from the strata need not have the same selection rates. A common situation is where the overpayment amount in a frame of claims is believed to be significantly correlated with the amount of the original payment to the physician or supplier. The frame may then be stratified into a number of distinct groups by the level of the original payment and separate simple random samples are drawn from each stratum. Separate estimates of overpayment are made for each stratum and the results combined to yield an overall projected overpayment.

The main object of stratification is to define the strata in a way that will reduce the margin of error in the estimate below which would be obtained by other sampling methods, as well as to obtain an unbiased estimate or an estimate with an acceptable bias. The standard literature, including that referenced in Appendix A, contains a number of different plans; the suitability of a particular method of stratification depends on the particular problem being reviewed, and the resources allotted to reviewing the problem. Additional discussion of stratified sampling is provided in Appendix B.

4. Cluster Sampling.--Involves drawing a random sample of clusters and reviewing everything or a sample of units in the sampled clusters. Unlike strata, clusters are groups of units that do not necessarily have strong similarities, but can be efficiently accessed for review purposes. For example, if the sampling unit is a beneficiary and the plan is to review each of the set of payments for each selected beneficiary, then the design is an example of cluster sampling with each beneficiary constituting a cluster of payments. The main point to remember (when sampling all the units in the cluster) is that the sample size for purposes of estimating the sampling error of the estimate is the number of clusters, not the total number of individual payments that are reviewed.

A challenge to the validity of a cluster sample that is sometimes made is that the number of sampling units in a cluster is too small. (A similar challenge to stratified sampling is also raised - i.e., that the number of sampling units in a stratum is too small). Such a challenge is usually misguided since the estimate of the total overpayment is a combination of the individual cluster (or, in the case of stratified sampling, stratum) estimates; therefore the overall sample size is important, but the

individual cluster (or stratum) sample sizes are usually not critical. Additional discussion of cluster sampling is provided in Appendix B.

Both stratification and cluster sampling are methods of grouping units. The former is frequently recommended when there is sufficient knowledge to group units that are similar in some aspect and potentially different from other units. The latter is frequently recommended when there are natural groupings that make a study more cost effective. When carried out according to the rules of probability sampling both of the methods, or a combination, are valid. The use of any of the methods described in this section will produce valid results when done properly.

5. Design Combinations.--A sample design may combine two or more of the methods discussed above. For example, clusters may be stratified before selection; systematic selection rather than simple random sampling may be used for selecting units within strata; or clusters may be subsampled using either simple random sampling or systematic sampling, to cite some of the possible combinations of techniques.

The benefits of stratification by claim amount may be achieved without actually stratifying if the frame is arranged in ascending order by the original payment amount and systematic sampling applied with a random start. That is because the systematic selection “balances out” the sample over the different levels of original payment in a manner similar to the effect of formal stratification. Thus systematic selection is often used in the hope that it will result in increased precision through “implicit stratification.”

B. Random Number Selection.--You must identify the source of the random numbers used to select the individual sampling units. Also document the program and its algorithm or table that is used; this documentation becomes part of the record of the sample and must be available for review. You must document any starting point if you are using a random number table or drawing a systematic sample. In addition, document the known seed value if a computer algorithm is used. You must document all steps in the random selection process exactly as you did them to ensure that the necessary information is captured for anyone attempting to replicate the sample selection.

There are a number of well-known, reputable software statistical packages (SPSS, SAS, etc.) and tables that may be used for generating a sample. One such package is RAT-STATS, available (at time of release of these instructions) through the Department of Health and Human Services, Office of Inspector General Web site. It is emphasized that the different packages offer a variety of programs for sample generation and do not all contain the same program services nor the same ease in operation. For any particular problem, your statistician or systems programmer should determine which package is best suited to the problem being reviewed.

C. Determining Sample Size.--The size of the sample (i.e., the number of sampling units) will have a direct bearing on the precision of the estimated overpayment, but it is not the only factor that influences precision. The standard error of the estimator also depends on (1) the underlying variation in the target population, (2) the particular sampling method that is employed such as simple random, stratified, or cluster sampling, and (3) the particular form of the estimator that is used (e.g., simple expansion of the sample total by dividing by the selection rate, or more complicated methods such as ratio estimation). It is neither possible nor desirable to specify a minimum sample size that applies to all situations. A determination of sample size may take into account many things, including the method of sample selection, the estimator of overpayment, and prior knowledge (based on experience) of the variability of the possible overpayments that may be contained in the total population of sampling units.

In addition to the above considerations, real-world economic constraints must be taken into account. As stated earlier, sampling is used when it is not administratively feasible to review every sampling unit in the target population. In practice, sample sizes may be determined by available resources. That does not mean, however, that the resulting estimate of overpayment is not valid as long as proper procedures for the execution of probability sampling have been followed. A challenge to the validity of the sample that is sometimes made is that the particular sample size is too small to yield meaningful results. Such a challenge is without merit as it fails to take into account all of the other factors that are involved in the sample design.

D. Documentation of Sampling Methodology.--You must provide complete documentation of the sampling methodology that you followed.

1. Documentation of Universe and Frame.--An explicit statement of how the universe is defined and elements included must be made in writing. Further, the form of the frame and specific details as to the period covered, definition of the sampling units, identifiers for the sampling units (e.g., claim numbers, carrier control numbers, etc.), and dates of service and source must be specified and recorded in your record of how the sampling was done. A record must be kept of the random numbers actually used in the sample and how they were selected. Sufficient documentation must be kept so that the sampling frame can be re-created, should the methodology be challenged. You must keep a copy of the frame.

2. Arrangement and Control Totals.--It is often convenient in frame preparation to array the universe elements by payment amount, e.g., low to high values, especially when stratification is used. At the same time, tabulate control totals for the numbers of elements and payment amounts.

3. Worksheets.--You must maintain documentation of the review and sampling process. All worksheets used by reviewers must contain sufficient information that allows for identification of the claim or item reviewed. Such information may include, for example:

- a. Name and identification number of the physician or supplier;
- b. Name and title of reviewer;
- c. The Health Insurance Claim Number (HICN), the unique claim identifier (e.g., the claim control number), and the line item identifier;
- d. Stratum and cluster identifiers, if applicable;
- e. The amount paid;
- f. The amount that should have been paid (either over or underpaid amount); and,
- g. The date(s) of service.

4. Overpayment/Underpayment Worksheets.--Worksheets should be used in calculating the net overpayment. The worksheet should include data on the claim number, line item, amount paid, audited value, amount overpaid, reason for disallowance, etc., so that each step in the overpayment calculation is clearly shown. Underpayments identified during reviews should be similarly documented.

E. Informational Copies to RO.--Send informational copies of your statistician-approved sampling methodology to the RO. The RO will keep the methodology on file and will forward to CO upon request. If this sampling methodology is applied routinely and repeatedly, you do not need to repeatedly send the methodology to the RO.

V. **Calculating the Estimated Overpayment.--**

A. The Point Estimate.--In simple random or systematic sampling the total overpayment in the frame may be estimated by calculating the mean overpayment, net of underpayment, in the sample and multiplying it by the number of units in the frame. In this estimation procedure, which is unbiased, the amount of overpayment dollars in the sample is expanded to yield an overpayment figure for the universe. The method is equivalent to dividing the total sample overpayment by the selection rate. The resulting estimated total is called the point estimate of the overpayment, i.e., the difference between what was paid and what should have been paid. In stratified sampling, an estimate is found for each stratum separately, and the weighted stratum estimates are added together to produce an overall point estimate.

In most situations the lower limit of a one-sided 90 percent confidence interval should be used as the amount of overpayment to be demanded for recovery from the physician or supplier. The details of the calculation of this lower limit involve subtracting some multiple of the estimated standard error from the point estimate, thus yielding a lower figure. This procedure, which, through confidence interval estimation, incorporates the uncertainty inherent in the sample design, is a conservative method that works to the financial advantage of the physician or supplier. That is, it yields a demand amount for recovery that is very likely less than the true amount of overpayment,

and it allows a reasonable recovery without requiring the tight precision that might be needed to support a demand for the point estimate. However, you are not precluded from demanding the point estimate where high precision has been achieved.

Other methods of obtaining the point estimate are discussed in the standard textbooks on sampling theory. Alternatives to the simple expansion method that make use of auxiliary variables include ratio and regression estimation. Under the appropriate conditions, ratio or regression methods can result in smaller margins of error than the simple expansion method. For example, if, as discussed earlier, it is believed that the overpayment for a sample unit is strongly correlated with the original paid amount, the ratio estimator may be efficient. The ratio estimator is the ratio of the sample net overpayment to the sample total original payment multiplied by the total of original paid dollars in the frame. If the actual correlation between the overpayment and the original paid amount is high enough, greater precision in estimation will be attained, i.e., the lower limit of the one-sided 90 percent confidence interval will be closer to the point estimate. Exercise caution about using alternatives such as ratio or regression estimation because serious biases can be introduced if sample sizes are very small. (The term bias is used here in a technical sense and does not imply a finding that treats the physician or supplier unfairly. A biased estimator is often used rather than an unbiased estimator because the advantage of its greater precision outweighs the tendency of the point estimate to be a bit high or low.)

B. Calculation of the Estimated Overpayment Amount.--The results of the sampling unit reviews is used to project an estimate of the overpayment amount. Each result shall be recorded except that a sampling unit's overpayment shall be set to zero if there is a limitation on liability determination made to waive physician or supplier liability for that sampling unit (per provisions found in §1879 of the Social Security Act (the Act)) and/or there is a determination that the physician or supplier is without fault as to that sampling unit overpayment (per provisions found in §1870 of the Act). Sampling units for which the requested records were not provided are to be treated as improper payments (i.e., as overpayments). Sampling units that are found to be underpayments, in whole or in part, are recorded as negative overpayments and are also used in calculating the estimated overpayment.

VI. Actions to be Performed Following Selection of Physician or Supplier and Sample.

NOTE: The instructions in this section dealing with notification and determination of location of the review do not supercede instructions for benefit integrity investigations, either planned or on-going.

A. Physician/Supplier Notification of Review and Review Site.--First, determine whether you will be notifying the physician or supplier of the review. Although in most cases prior notification is provided, the physician or supplier is not always notified before the start of the review. When not giving advance notice, you must obtain RO advance approval as required by applicable instructions. When giving advance notice, provide written notification by certified mail with return receipt requested (retain all receipts).

Second, regardless of whether you give advance notice or not, determine where you will conduct your review, at the physician's or supplier's site(s) or at your offices (contractor site).

1. Written Notification.--Include in the notification an explanation of why the review is being conducted (i.e., why the physician or supplier was selected and the period of review), the list of claims that require medical records or other supporting documentation, where the review will take place (physician/supplier site or contractor site), information on appeal rights, and an explanation of the possible methods of monetary recovery if claims are denied upon review. Also include an explanation of how results will be projected to the universe.

When advance notification is given, physicians and suppliers have 30 calendar days to submit (for contractor site reviews) or make available (for physician/supplier site reviews) the requested documentation. Advise the physician or supplier that should requested documentation not be submitted or made available by the end of 30 calendar days, you will start the review and you will deny those claims for which there is no documentation. You do not have to request all

documentation at time of notification of review. For example, you may decide to request one-half of the documentation before you arrive, and then request the other half following your arrival at the physician/supplier's site. The time limit for submission or production of requested documentation may be extended at your discretion.

When advance notification is **not** given, give the physician or supplier the written notification when you arrive at their site.

2. Determining Review Site.--

a. Physician/Supplier Site Reviews.--Physician/supplier site reviews are performed at the physician's or supplier's location(s). Considerations in determining whether to conduct a physician/supplier site review include:

- The extent of aberrant patterns identified;
- The presence of multiple program integrity issues;
- Evidence or likelihood of fraud; and/or,
- Past failure(s) of a physician or supplier to submit requested
- Medical records in a timely manner or as requested.

b. Contractor Site Reviews.--Contractor site reviews are performed at the contractor's location.

B. Meetings to Start and End the Review.--In-person meetings to start and end the review are encouraged, but are not required or always feasible. If you hold an in-person meeting at the start of the review, explain the scope and purpose of the review and discuss the next steps at the end of the review. Attempt to answer the physician's or supplier's questions related to the review.

During an exit meeting, you may discuss the basic or preliminary findings of the review. Give the physician or supplier an opportunity to discuss or comment on the claims decisions that were made. Advise the physician or supplier that a demand letter detailing the results of the review and the statistical sampling will be sent if an overpayment is determined to exist.

C. Conducting the Review.--Following your receipt of the requested documentation (or the end of the period to submit or make available the requested documentation, whichever comes first), start your review of the claims. Obtain additional documentation as necessary for an objective and thorough evaluation of the payments that have been made. Use physician consultants and health professionals in the various specialties as necessary to review or approve decisions involving medical judgment. The review decision is made on the basis of the Medicare law, CMS rulings, regulations, national coverage determinations, and regional/local carrier medical review policies that were in effect at the time the item(s) or service(s) was provided.

Document all findings made so that it is apparent from your written documentation if your initial determination has been reversed. Document the amount of all overpayments and underpayments and how they were determined.

You are encouraged to complete your review and calculate the net overpayment within 90 calendar days of the start of the review (i.e., within 90 calendar days after you have either received the requested documentation or the time to submit or make available the records has passed, whichever comes first). However, there may be extenuating circumstances or circumstances out of your control where you may not be able to complete the review within this time period (e.g., you have made a fraud referral to the OIG and are awaiting their response before pursuing an overpayment).

Your documentation of overpayment and underpayment determinations must be clear and concise. Include copies of the local medical review policy and any applicable references needed to support individual case determinations. Compliance with these requirements facilitates adherence to the physician and supplier notification requirements.

VII. Overpayment Recovery.--

A. Recovery from Physician or Supplier.--Once an overpayment has been determined to exist, proceed with recovery based on applicable instructions (see MCM §7130). Regardless of which unit within the carrier pursues the overpayment, include in the overpayment demand letter information about the review and statistical sampling methodology followed.

An explanation of the sampling methodology that was followed should include:

- A description of the universe, the frame, and the sample design;
- A definition of the sampling unit, the sample selection procedure, and the numbers and definitions of the strata and size of the sample, including allocations, if stratified;
- The time period under review;
- The sample results, including the overpayment estimation methodology and the calculated sampling error as estimated from the sample results; and
- The amount of the actual overpayment/underpayment from each of the claims reviewed.

Also include a list of any problems/issues identified during the review, and any recommended corrective actions.

B. Informational Copy to RO.--Send an informational copy of the demand letter to the RO. The RO will maintain copies of demand letters and will forward to CO upon request. If the demand letter is used routinely and repeatedly, you do not need to repeatedly send it to the RO.

VIII. Corrective Actions.--Take other corrective actions you deem necessary (such as payment suspension, imposition of civil money penalties, institution of pre- or post-payment review, additional edits, etc.)

IX. Changes Resulting from Appeals.--If the decision issued on appeal contains either a finding that the sampling methodology was not valid, and/or reverses the revised initial claim determination, you must take appropriate action to adjust the extrapolation of overpayment.

A. Sampling Methodology.--If the decision issued on appeal contains a finding that the sampling methodology was not valid, there are several options for revising the estimated overpayment based upon the appellate decision:

1. If the decision issued on appeal permits correction of errors in the sampling methodology, you must revise the overpayment determination after making the corrections. Consult with CO through RO to determine whether such an action is consistent with the hearing officer (HO), administrative law judge (ALJ) or Departmental Appeals Board (DAB) decision, or court order.

2. You may elect to recover the actual overpayment related to the sample claims and then initiate a new review of the physician or supplier. If the actual overpayments related to the sampling units in the original review have been recovered, then these units should be eliminated from the sampling frame used for any new review. Consult with CO through RO to determine whether such an action is consistent with the HO, ALJ or DAB decision, or court order.

3. You may conduct a new review (using a new methodology) for the same time period as was covered by the previous review. Before employing this option, consult with CO through RO to verify that the action is consistent with the HO, ALJ or DAB decision, or court order. If this option is chosen, you may not recover the overpayments on any of the sample claims found to be in error in the original sample.

B. Revised Initial Determination--If the decision on appeal upholds the sampling methodology but reverses one or more of the revised initial claim determinations, the estimate of overpayment must be recomputed.

Resources:

American Institute of Certified Public Accountants, Statistical Sampling Subcommittee, *Audit Sampling*, 1999.

Arkin, H., *Handbook of Sampling for Auditing and Accounting*, 1984.

Cochran, W. G., *Sampling Techniques*, 3rd ed., New York: John Wiley and Sons, 1977.

Deming, W. E., *Sample Design in Business Research*, New York: John Wiley and Sons, 1960 (Paperback 1990).

Hansen, M. H., Hurwitz, W. W., and Madow, W. G., *Sample Survey Methods and Theory*, New York: John Wiley and Sons, 1953 (Paperback 1993).

Hedayat, A., Bekas, K. S., *Design and Inference in Finite Population Sampling*, John Wiley & Sons, New York, 1991.

Kish, L., *Survey Sampling*, New York: John Wiley and Sons, 1967, 2nd printing. (Paperback 1995).

Levy, P. and Lemeshow, S., *Sampling of Populations Methods and Applications*, 3rd ed., John Wiley & Sons, 1999.

Scheaffer, R. L., Mendenhall, W., and Ott, L., *Elementary Survey Sampling*, 5th ed., Duxbury Press, 1996.

Som, R. K., *Practical Sampling Techniques*, M. Dekker, New York, 1996, 2nd ed.

Additional Discussion of Stratified Sampling

Generally, one defines strata to make them as internally homogeneous as possible with respect to overpayment amounts, which is equivalent to making the mean overpayments for different strata as different as possible. Typically, a proportionately stratified design with a given total sample size will yield an estimate that is more precise than a simple random sample of the same size without stratifying. The one highly unusual exception is one where the variability from stratum mean to stratum mean is small relative to the average variability within each stratum. In this case, the precision would likely be reduced, but the result would be valid. It is extremely unlikely, however, that such a situation would ever occur in practice. Stratifying on a variable that is a reasonable surrogate for an overpayment can do no harm, and may greatly improve the precision of the estimated overpayment over simple random sampling. While it is a good idea to stratify whenever there is a reasonable basis for grouping the sampling units, failure to stratify does not invalidate the sample, nor does it bias the results.

If it is believed that the amount of overpayment is correlated with the amount of the original payment and the universe distribution of paid amounts is skewed to the right, i.e., with a set of extremely high values, it may be advantageous to define a “certainty stratum”, selecting all of the sampling units starting with the largest value and working backward to the left of the distribution. When a stratum is sampled with certainty, i.e., auditing all of the sample units contained therein, the contribution of that stratum to the overall sampling error is zero. In that manner, extremely large overpayments in the sample are prevented from causing poor precision in estimation. In practice, the decision of whether or not to sample the right tail with certainty depends on fairly accurate prior knowledge of the distribution of overpayments, and also on the ability to totally audit one stratum while having sufficient resources left over to sample from each of the remaining strata.

Stratification works best if one has sufficient information on particular subgroups in the population to form reasonable strata. In addition to improving precision there are a number of reasons to stratify, e.g., ensuring that particular types of claims, line items or coding types are sampled, gaining information about overpayments for a particular type of service as well as an overall estimate, and assuring that certain rarely occurring types of services are represented. Not all stratifications will improve precision, but such stratifications may be advantageous and are valid.

Given the definition of a set of strata, the designer of the sample must decide how to allocate a sample of a certain total size to the individual strata. In other words, how much of the sample should be selected from Stratum 1, how much from Stratum 2, etc.? As shown in the standard textbooks, there is a method of “optimal allocation,” i.e., one designed to maximize the precision of the estimated potential overpayment, assuming that one has a good idea of the values of the variances within each of the strata. Absent that kind of prior knowledge, however, a safe approach is to allocate proportionately. That is, the total sample is divided up into individual stratum samples so that, as nearly as possible, the stratum sample sizes are in a fixed proportion to the sizes of the individual stratum frames. It is emphasized, however, that even if the allocation is not optimal, using stratification with simple random sampling within each stratum does not introduce bias, and in almost all circumstances proportionate allocation will reduce the sampling error over that for an unstratified simple random sample.

Additional Discussion of Cluster Sampling

Selecting payments in clusters rather than individually usually leads to a reduction in the precision of estimation. However, your reasons for using cluster sampling instead of simple random sampling may be driven by necessity and/or cost-savings related to the location of records or the nature of a record. For example, for medical review to determine the appropriateness of certain charges for a beneficiary it may be necessary to examine the complete medical record of the patient. This then may allow for review of claims for several services falling within the selected review period. In another instance, the medical records that you must review may be physically located in a cluster (e.g., the same warehouse, the same file drawer, the same folder) with the medical records for other similar claims and it is cost effective to select units from the same location. Whenever the cost in time and other resources of selecting and auditing clusters is the same as the cost of simple random

sampling the same number of payments, it is better to use simple random sampling because greater precision will be attained.

When reviewing all the units in each cluster, the sample size is the number of clusters, not the number of units reviewed. This is single-stage cluster sampling, a method frequently used when sampling beneficiaries. One may choose to review a sample of units within each cluster rather than all units. Textbooks that cover the topic of multi-stage sampling provide formulas for estimating the precision of such sample designs. One example for which multi-stage sampling might be an appropriate choice of design is the case of reviewing a supplier chain where records are spread out among many locations. The first-stage selection would be a sample of locations. At the second stage a subsample of records would be selected from each sampled location.