

therefore only the actual overpayment could be recovered. Dec. I at 35, filed in the record as Exh. 4 at 38.¹

The Council subsequently remanded the case on the grounds that the Centers for Medicare & Medicaid Services (CMS) and/or its contractors, including AdvanceMed, the Program Safeguard Contractor (PSC) that performed the audit in this case, had not been afforded an opportunity to participate in the ALJ hearing. See Council Order of Remand in *John Sanders, MD*, Docket No. M-10-564, issued March 26, 2010. On remand, the ALJ provided the PSC with an opportunity to testify, and incorporated the earlier testimony of the appellant's expert statistician, and the independent expert statistician retained by the Office of Medicare Hearings and Appeals (OMHA) into the record, by agreement. The ALJ then issued a new decision, dated December 14, 2010, determining that "on the whole, . . . the sampling methodology utilized by AdvanceMed did not comply with Medicare requirements, and the extrapolated overpayment calculation was invalid." Dec. II at 9.

The appellant did not request Council review of the second ALJ decision. However, CMS filed a referral memorandum, seeking own motion review by the Council. The CMS memorandum is entered into the record as Exhibit MAC-1. The CMS memorandum contests the ALJ's determination that the sampling methodology did not comply with Medicare requirements and therefore the extrapolated overpayment calculation was invalid. Exh. MAC-1. The appellant has filed a timely response to the CMS memorandum, which is entered into the record as Exh. MAC-2.

The Council has carefully considered the record before the ALJ, as well as the CMS memorandum and the appellant's response. The Council hereby modifies the ALJ's second decision, concurring in the determination that the sampling was sufficiently flawed to preclude calculation of an overpayment by extrapolation. However, the Council's reasons for this determination differ from the reasons set forth in the ALJ's second decision. Specifically, as explained in the Council's Analysis below, many of the reasons cited in the Analysis section of the ALJ's second decision do not provide a basis for concluding that the sample was invalid.

¹ The ALJ decision issued November 27, 2009 (which was earlier vacated by the Council), is referred to herein as Dec. I. The ALJ decision issued December 14, 2010 (after the Council's remand), is referred to as Dec. II.

However, there were two sampling issues raised by the appellant, which were neither sufficiently explained nor corrected by the PSC, and which were the subject of testimony from the independent expert statistician. These two errors are not addressed in or rebutted by the CMS memorandum. The errors are: 1) the PSC provided the independent statistical expert with sample data which assigned some claims to the wrong stratum; and 2) the PSC provided the independent expert with a second CD containing an Excel set of sample data with significant discrepancies from the first set of data, and the PSC was unable to clarify the discrepancies, to identify which set of data was applicable, or to explain the significance of the second set of data. The Council finds that these errors and inconsistencies in the original sampling preclude use of the sample to extrapolate an overpayment to the full universe of claims. Therefore, the Council modifies the reasoning that underlies the ALJ's conclusion in his second decision, but concurs in the conclusion that the results of the sampling cannot be used to extrapolate an overpayment amount. The appellant remains financially liable for the overpayments shown in the individual claims within the sample, but is not financially liable for any extrapolated amount.

BACKGROUND

Below, the Council sets out a brief synopsis of the pertinent background and procedural history of this case. Further information can be found in the Council's earlier remand order (*John Sanders, MD*, Docket No. M-10-564, issued March 26, 2010).

AdvanceMed Corporation, a Program Safeguard Contractor (PSC), conducted a post-payment review of claims submitted to Medicare for evaluation and management (E&M) services furnished by the appellant physician, with dates of service between January 1, 2004 and December 31, 2005. Exh. 1 at 151-52. In a sample of approximately sixty claims reviewed, the PSC denied or downcoded 227 of the 235 line items (each consisting of one HCPCS code). *Id.* at 165.² Based on this review and on an extrapolation from the statistical sampling, CIGNA, the Medicare Administrative

² The Centers for Medicare and Medicaid Services (CMS) has established uniform national definitions of services, codes to represent services, and payment modifiers to the codes. 42 C.F.R. § 414.40(a). The Medicare coding system, Healthcare Common Procedural Coding System (HCPCS) is based on the American Medical Association's (AMA's) Physician's Current Procedural Terminology (CPT).

Contractor (MAC), requested a return of \$211,218 that it claimed was paid to the appellant in error. *Id.* at 138. Upon redetermination, the contractor upheld the overpayment determination. *Id.* at 130-133.

The appellant requested reconsideration, and advanced contentions about both the findings of overpayments in individual claims and the methodology used in the statistical sampling and extrapolation of the overpayment amount. Exh. 1 at 115-29. The Qualified Independent Contractor (QIC) upheld a majority of the fully or partially denied claims in the sample for generally the same reasons as the contractor. *Id.* at 73-102. However, the QIC did adjust payment upwards for approximately eighteen of the claims based on the appellant's contentions. *Id.* The QIC also found the appellant liable for the extrapolated overpayment, rejecting the appellant's contentions as to the invalidity of the sampling and extrapolation methodologies. *Id.* at 95-96.

The appellant requested an ALJ hearing on March 18, 2008. Exh. 1 at 7-15. The ALJ held a prehearing conference, and a hearing on June 9, 2009. Prehearing Conference CD (June 20, 2008); ALJ Hearing CD (June 9, 2009). The ALJ heard testimony from the appellant on services and billing for the claims in the sample, and from an independent statistical expert and the appellant's statistical expert on the methods used in the sampling and extrapolation. ALJ Hearing CD (June 9, 2009) at approximately 11:43 a.m. to 12:01 p.m., and 1:00 to 3:00 p.m.

On November 27, 2009, the ALJ issued a partially favorable decision in which he determined some services in the sample (in parts of approximately 35 claims) were reasonable and necessary and met coverage guidelines. *Id.* at 15-33. For services that remained denied, the ALJ found the appellant liable for the overpayment pursuant to section 1879 of the Social Security Act (Act) and not entitled to waiver under section 1870 of the Act. *Id.* at 33-34.

With regard to the statistical sampling used to calculate the overpayment, the ALJ noted that while CMS contractors may use statistical sampling to calculate overpayments, the "sampling study must be based upon appropriate sampling and computed by valid statistical methods to establish *prima facie* evidence of the number and amount of claims or requests for payment." *Id.* at 15. In this case, both the appellant's statistical expert and the independent statistical expert, "testified as to flaws

in the sampling size and the stratification of the sample utilized by PSC." *Id.* Based on "unanimous expert testimony," the ALJ determined that "the sampling methodology utilized by the PSC did not comply with Medicare requirements, and therefore the extrapolated overpayment calculation was invalid." *Id.* at 15. The ALJ concluded that the appellant remained liable for a number of services in the sample that were not payable by Medicare. *Id.* at 35. The ALJ further concluded also that the statistical sampling methodology used to calculate the extrapolated overpayment was invalid and therefore only the actual overpayment could be recovered. *Id.*

CMS requested own motion review of the ALJ's decision by the Council, on the ground that the record did not show that potential participants, including the PSC, were afforded notice and an opportunity to participate in the hearing. Exh. MAC-1 at 1-11, *citing* 42 C.F.R. §§ 405.1010, 405.1020, and 405.1022.³

On March 26, 2010, the Council remanded the case to the ALJ to provide an opportunity for the PSC to participate in a hearing, and to make a complete record of the evidence. The Council noted that Medicare regulations provide that CMS and/or one or more of its contractors may elect to participate in the hearing process. 45 C.F.R. § 405.1010(a). Because the PSC had performed the sampling and had provided a copy of its data to the other statisticians to review, it was both fundamentally fair and required by regulation that it be given an opportunity to participate in the ALJ hearing when the validity of the sampling was at issue.⁴

³ In its referral memorandum, dated January 20, 2010, CMS also contended that the types of concerns raised by both statistical experts in the case should not have served as a basis for invalidating the extrapolation of the overpayment. Exh. 1 at 6-7.

⁴ The Council disagrees with the ALJ's characterization of its remand as based on an "expansive" interpretation of 42 C.F.R. § 405.1020. Precisely because § 405.1010(a) allows CMS and/or its contractor(s) to elect to participate in the hearing process, those entities must receive notice under § 405.1020 in order to exercise their right to participate. In fact, subsection 405.1020(c)(2) makes explicit reference to the fact that the notice of hearing requires "all parties to the ALJ hearing (and any potential participant from CMS or its contractor who wishes to attend the hearing)" to reply to the hearing notice. (Emphasis added.) In this case, where the validity of the PSC's sampling and extrapolation was at issue, and where the ALJ invalidated the sampling and extrapolation for the first time during the appeal process, it was particularly important for the PSC to be afforded the opportunity to participate in the hearing. On remand, the PSC did elect to participate.

On remand, after a prehearing conference, the ALJ held a hearing and took testimony from a representative of the PSC, Mr. Landtroop. The earlier testimony of the two statistical experts, Dr. Haller and Dr. Rhode, was incorporated into the record, by agreement.

In accordance with the Council's remand order, the ALJ issued a new decision. That decision reviewed the background of the case, the earlier testimony of the two statistical experts, and Mr. Landtroop's testimony on behalf of the PSC. Dec. II at 1-3. Based on the testimony, the ALJ found that because certain parts of the sampling methodology utilized by the PSC did not comply with Medicare requirements, the extrapolated overpayment calculation was invalid and the appellant was not required to reimburse the claimed extrapolated amount. *Id.* at 9. As explained above, this is the ALJ decision that CMS seeks to challenge in its request for own motion review. Six parts of the record are of particular relevance to deciding the issues presented in the CMS request and the appellant's response. These parts of the record are each summarized below.

SUMMARY OF ISSUES AND POSITIONS

Testimony of Appellant's Statistics Expert

During the hearing, the appellant's statistical expert, Dr. Rhode, testified that in a number of respects he did not have enough information to opine on whether the statistical sampling was valid. ALJ Hearing CD (June 9, 2009) at 11:45 to 11:57 a.m. He testified that he lacked information about why the sampling unit was the claim rather than the billing line. He also lacked information on the distribution of the sample (i.e., whether it was normal (bell-shaped) or skewed in some manner). *Id.* He further opined that since fully denied claims had been omitted from the sampling and it was not clear how the overpayment rate was defined, such factor could have affected the numbers used to compute the overpayment projection. *Id.*

Testimony of Independent Statistical Expert

The statistical expert retained by the ALJ, Dr. Haller, prepared a written report, or "case review," submitted prior to the hearing. Exh. 3 at 1-6. Initially, Dr. Haller reported that the first Excel CD file he received from the PSC (entitled MedicalReviewSpreadSheetwithoverpayment.xls) had what appeared to be data from the audit of the random sample defined in PSC

former statistician Dr. Moody's (Dr. Moody's) memo of January 30, 2007 (Exh. 1 at 171-175). Exh. 3 at 1-6. However, when Dr. Haller sorted the data, there were 23 beneficiaries for Stratum 1 (amounts paid < \$350) and 29 beneficiaries for Stratum 2 (amounts paid > or = \$350), and some of the beneficiaries were identified as being in the wrong stratum. *Id.* at 2. This was inconsistent with Dr. Moody's plan for the sampling, which envisioned 30 beneficiaries in each stratum. Exh. 1 at 171-72.

Dr. Haller explained that a second Excel CD file (entitled Samplecollapse.xls) was furnished by the PSC for his review. Exh. 3 at 2. This second file contained data on amounts paid and overpaid for 30 beneficiaries in each of the two strata. *Id.* However, Dr. Haller stated that he could not explain the differences between the two Excel files. *Id.* at 2-3. His report concludes by asking that AdvanceMed clarify the discrepancy between the data in the two files. *Id.* at 4. The PSC has been unable to do so, as Dr. Moody is no longer with the PSC.

In his written testimony, Dr. Haller also explained that based on the use of an Optimum Allocation strategy, he believed the sample size should have been 34 beneficiaries for Stratum 1 and 55 beneficiaries for Stratum 2. Exh. 3 at 3. However, acknowledging the language in the Medicare Program Integrity Manual that discourages challenges to sample size alone (MPIM, ch. 3, § 3.10.4.3.), he referred to the sample size as "adequate if other conditions to be discussed below are met." Exh. 3 at 3.

Then Dr. Haller explained that given the 90% two-sided confidence interval, and the 15.7% precision, it is important to examine the *distribution* of the determined overpayments. Exh. 4 at 3-4. Based on his analysis, he identified the determined overpayments as skewed to the right, and therefore he used a "natural logarithm transformation" to render the data in a normal distribution before he calculated the confidence interval for the mean overpayment to the universe (point estimate). *Id.* at 4. Based on this procedure, he calculated a 95% probability that the total overpayments to the appellant were at least \$145,740, rather than the \$211,218 amount computed by AdvanceMed. *Id.*

Testimony of PSC Representative on Statistics Issues

During the hearing on remand (on November 3, 2010), Mr. Landtroop, who holds a Masters degree and now serves as chief statistician for AdvanceMed, testified about the statistical issues in the case. ALJ Hearing CD (Nov. 3, 2010) at 10:11 to 10:32 a.m. Mr. Landtroop had reviewed the written testimony from Dr. Haller and the summary of Dr. Rhode's testimony contained in the ALJ's written decision. *Id.* at 10:23 a.m.; see also Dec. I. at 3, 18, 35. Dr. Moody, who had supervised the audit in 2006 to 2007, was no longer employed by AdvanceMed. *Id.*

Mr. Landtroop stated first that although the confidence interval of the AdvanceMed data was only to 15% precision, and Dr. Moody's memo planned for 10% precision, there is no requirement in Medicare policy guidelines which require a precision percentage of 10% or less. ALJ Hearing CD (Nov. 3, 2010) at 10:13 to 10:17 a.m. He also stated that although he had read Dr. Haller's testimony about the mix-up on the files that were sent, he "really couldn't speak to that, really couldn't say what had happened." *Id.* at 10:17-10:18 and 10:30 a.m.

Mr. Landtroop addressed Dr. Haller's testimony about the sampled overpayments being skewed to the right in two ways. ALJ Hearing CD (Nov. 3, 2010) at 10:20 to 10:23 a.m. First, he asserted that simply saying the overpayments are not distributed normally is not enough to invalidate the extrapolation. *Id.* at 10:20 a.m.. According to Mr. Landtroop, given the statistical assumptions of the Central Limit Theorem, one would have to show that the distribution of the sample results was abnormal in comparison to the distribution of the results in the sampling frame. *Id.* He submits that it would not be realistic to perform such a procedure, and that the sample of 60 claims was more than enough to ensure the sample average was normally distributed. *Id.*

Second, Mr. Landtroop questioned Dr. Haller's use of a "natural logarithm transformation" to revise the data before doing an extrapolation. ALJ Hearing CD (Nov. 3, 2010) at 10:21 to 10:22 a.m. Because four or five of the sample claim overpayments were zeros or negative numbers (representing no overpayment or an underpayment), it would not be possible to use a natural logarithm process because zeros and negative numbers have no defined natural logarithms. *Id.*

Second ALJ Decision

In the decision issued December 14, 2010, the ALJ concluded that AdvanceMed's sampling methodology did not comply with Medicare requirements, the extrapolated overpayment calculation was invalid, and the appellant is not required to reimburse the assessed extrapolated amount. Dec. II at 9. The ALJ identified the following reasons for his conclusions:

- Dr. Haller stated that the confidence interval of the AdvanceMed data was only to 15% precision, and he estimated a correct extrapolation amount that was nearly one-third less than the AdvanceMed result.
- Dr. Rhode stated a variety of doubts about the AdvanceMed sampling.
- While Mr. Landtroop argued that the 10% precision is not a strict requirement, the Manual provides, "In most situations the lower limit of a one-sided 90 percent confidence interval shall be used as the amount of overpayment to be demanded for recovery from the provider or supplier."
- Mr. Landtroop did not agree with the recalculated extrapolation estimate of Dr. Haller. However, he did concede that the Central Limit Theorem assumes a normal distribution of results, and he did not challenge the assertion that the results in this case were skewed.

Dec. II at 9.

CMS Memorandum

In its request for own motion review by the Council, CMS asserts that there is an error of law material to the outcome of the claim, and that the ALJ's decision is not supported by a preponderance of the evidence. Exh. MAC-1 at 1. Overall, the CMS memo contends that none of deficiencies Dr. Haller and Dr. Rhode identified in the sampling and extrapolation methods are sufficient to invalidate the sampling and extrapolation. *Id.* at 2.

More specifically, the CMS memorandum defends the PSC's use of data with a 15.69% precision estimate, by explaining that the

Medicare Program Integrity Manual contemplates the use of the lower bound of a confidence interval to increase the probability that the overpayment demand is equal to or less than the actual overpayment. Exh. MAC-1 at 5-7, *citing*, Pub. 100-08, MPIM, ch. 3, §§ 3.10.2, 3.10.5.1.

The CMS memorandum also takes issue with the ALJ's reliance, in part, on Dr. Haller's testimony that the skewed distribution of the sampled overpayment results raises questions about the extrapolation, and with Dr. Haller's use of alternate computations to arrive at an extrapolation of \$145,740. Referring to Mr. Landtroop's testimony, the CMS memo points out that the fundamental assumption of the Central Limit Theorem is that the *sample averages* will be normally distributed. The fact that the overpayment sampling units in one sample may be skewed does not mean that the sample averages, if successive samples were to be taken, are not normally distributed. Parsing Dr. Haller's testimony further, CMS also notes that Dr. Haller did not say that the PSC calculation of the extrapolation was invalid; rather, he testified that he had found "a better way to approach the problem," that resulted in "a more accurate estimate of the total overpayment to the universe." Exh. MAC-1 at 8, *citing* Exh. 3 at 4. CMS asserts that pursuant to CMS Ruling 86-1, the appellant has the burden of demonstrating that the sampling methodology used was invalid; it is not the contractor's responsibility to establish why it did not use a different (or more precise) design. Exh. MAC-1 at 8. To the extent that the ALJ relied on Dr. Haller's approach and calculation as a basis for finding the PSC's methodology invalid, CMS submits that his decision is erroneous.

Appellant's Response

In response to the CMS memorandum, the appellant asserts that the evidence the ALJ relied on demonstrates that the extrapolated overpayment amount was too high to be accurate. Exh. MAC-2 at 1-3. If this extrapolated (or estimated) overpayment amount is shown to be too high, then the appellant contends that the CMS memorandum and the CMS Ruling 86-1 state that the appellant has met its burden, and the extrapolation should be invalidated. *Id.*, *citing* Exh. MAC-1 at 2.

The appellant also contends that there is a preponderance of evidence supporting the ALJ's conclusion that the sampling and extrapolation were invalid. Exh. MAC-2 at 2. In appellant's view, this evidence includes the fact that the initial data

supplied to Dr. Haller were erroneous and inconsistent; the fact that the sample size used did not comport with the Optimum Allocation theory; and the significantly lower extrapolation arrived at when Dr. Haller corrected for skewing in the data; *inter alia. Id.*

LEGAL AUTHORITY

CMS Ruling 86-1 describes the agency's policy on the use of statistical sampling to project overpayments to Medicare providers and suppliers. The Ruling also outlines the history and authority, both statutory and precedential, for the use of statistical sampling and extrapolation by CMS in calculating overpayments. The Council incorporates that discussion by reference here. The Ruling provides, in part:

Sampling does not deprive a provider of its rights to challenge the sample, nor of its rights to procedural due process. Sampling only creates a presumption of validity as to the amount of an overpayment which may be used as the basis for recoupment. The burden then shifts to the provider to take the next step. The provider could attack the statistical validity of the sample, or it could challenge the correctness of the determination in specific cases identified by the sample (including waiver of liability where medical necessity or custodial care is at issue). In either case, the provider is given a full opportunity to demonstrate that the overpayment determination is wrong. If certain individual cases within the sample are determined to be decided erroneously, the amount of overpayment projected to the universe of claims can be modified. If the statistical basis upon which the projection was based is successfully challenged, the overpayment determination can be corrected.

CMS Ruling 86-1 at 86-1-9, 86-1-10.

The Medicare Program Integrity Manual provides guidance to contractors in conducting statistical sampling for use in estimating overpayment amounts. The instructions are intended to ensure that a statistically valid sample is drawn and that statistically valid methods are used to project overpayments where a review of claims indicates that overpayments have been made. The MPIM describes the purpose of its guidance as follows:

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

MPIM, ch. 3, § 3.10.1.1.

The MPIM further provides that a contractor may employ any sampling methodology that results in a "probability sample," and defines the requirements for a valid probability sample. See MPIM, ch. 3, § 3.10.2. The Manual then states:

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.

MPIM, ch. 3, § 3.10.2. The MPIM recognizes that a number of sampling designs are acceptable, including: simple random sampling, systematic sampling, stratified sampling, and cluster sampling, or a combination of these. MPIM, ch. 3, § 3.10.4.1.

The MPIM provides the following guidance with respect to selecting the 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 shall 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 determining the sample size to be used, the PSC or ZPIC BI unit or the contractor MR unit shall also consider their available resources. That does not mean, however, that the resulting estimate of overpayment is not valid, so 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.

MPIM, ch. 3, § 3.10.4.3.

ANALYSIS

The Council concurs with the ALJ's conclusion that the sampling conducted in this case was substantively flawed, and therefore the Council limits the overpayment recovery to the sum of the actual sampled claims without extrapolation. However, the problem with the sampling is not the general methodology design issues identified by the ALJ in his Analysis and contested in the CMS memorandum. In fact, the Council agrees with a number

of the points made in the CMS memorandum, for the reasons explained below. The Council, however, finds that the actual performance of the sampling contained sufficient flaws and unanswered questions as to render any extrapolation in this case subject to substantial errors and inequities to the appellant.

Problems in the Assignment of Claims to Strata

There are two major, related shortcomings in the sampling here, which cannot be corrected at this juncture. Therefore, notwithstanding the significant resources that the PSC has expended in conducting this audit, and the detailed and thorough planning and design methodology for the audit thoroughly recorded in Dr. Moody's January 30, 2007 memo, the Council must limit the overpayment demand to the total sum of the individual overpayment amounts identified in the cases sampled and reviewed by the QIC and the ALJ.

The first of these two flaws is that either the samples themselves were not drawn correctly or the claims were not correctly assigned to the correct stratum in every case, consistent with the probability sample design. The use of one or more probability samples in calculating overpayments is premised, *inter alia*, on the accuracy and representativeness of the sample or samples in representing the stratum from which they are drawn. That is why, in explaining the statistical sampling procedures, the Medicare Benefits Policy Manual states:

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.

MPIM, ch. 3, § 3.10.2 (emphasis added).

However, in this case, the probability sample design set forth in Dr. Moody's January 30, 2007 memorandum was not properly executed. The design called for a stratified random sampling, with 30 cases in Stratum 1 (payment of < \$350) and 30 cases in Stratum 2 (payment = or > \$350). Exh. 1 at 171-175. This design was based on Dr. Moody's analysis of the characteristics of the universe to be sampled. *Id.* However, when Dr. Haller opened and sorted the first Excel file of audit data that the

PSC provided for his review, he found only 23 beneficiaries in Stratum 1 and 29 beneficiaries in Stratum 2. Exh. 3 at 2. Moreover, Dr. Haller found that the PSC had placed some beneficiaries in Stratum 2 for whom the total amount paid the provider was less than \$350, which was inconsistent with the definition of Stratum 2. *Id.*

These errors are significant for a series of reasons. If claims in the sample were assigned to incorrect strata, it is possible, if not likely, that claims in the frame were also assigned to incorrect strata, but it is not possible to know at this juncture how widespread these errors were. If claims in the sample that were assigned to the wrong strata were in fact used in calculating key variables, such as error rates, average overpayments, point estimate(s), upper and lower confidence bounds, and precision estimate(s), then those variables and results would be inaccurate. Extrapolation might well have the effect of multiplying errors in this process.

At this point, it is not possible to perform a precise assessment of the nature and extent of any such errors in the way the sample was drawn. It is also not possible to identify and correct the errors without starting the entire audit process over, i.e., re-assigning every claim listed in the frame to assure that it is assigned to the correct stratum and then drawing a new random sample for review.

The second major concern the Council has with the sampling process, as raised by the appellant before the ALJ, relates to the uncertainty and inconsistency of the data recorded in two different and unidentified Excel CD files. Dr. Haller stated that the PSC submitted a second Excel file for his review. Exh. 3 at 2. This second file, according to Dr. Haller, did contain data on amounts paid and overpaid for 30 beneficiaries in each of the two strata, 1 and 2. *Id.* However, Dr. Haller stated that he could not explain the apparent difference between these two files, and he asked the PSC (in 2009) to clarify the discrepancy between the two files representing the probability sample. *Id.* at 2-4. The PSC did not respond to this request. At the hearing, the PSC's representative, Mr. Landtroop, stated:

And you're right, your Honor, there does appear to be this confusion about the Excel file that Dr. Haller received I can't speak to that. I wasn't an employee here at the time that occurred.

ALJ Hearing CD (Nov. 3, 2010) at 10:30 a.m. No one has been able to explain which file represents the correct data, and none of the statistical experts nor the Council has been able to resolve the substantial discrepancy between the two sets of data.

Unfortunately, the submission of data to Dr. Haller on the second CD raises additional questions about the accuracy of the audit. The first CD had data on 52 sampled claims, the second CD had data on 60 sampled claims. Exh. 3 at 2. Thus, the second CD contained a probability sample which is different (in part or in whole) from the probability sample contained in the first CD. Dr. Haller did not become involved in the case until approximately two and one-half years after the medical reviews were performed, so he is unable to ascertain when the second CD was compiled. Not only does the second CD not answer the questions raised by the data errors on the first CD, it adds a number of additional questions and additional room for inaccuracies in the sampling and extrapolation process. Provisions in the MPIM requiring detailed documentation of the sampling are intended to prevent problems of this type. See MPIM, ch. 3, §§ 3.10.4.4 (requiring documentation of the sampling methodology); 3.10.4.4.1. (requiring documentation of the universe and frame); and 3.10.4.4.3. (requiring documentation of the review and sampling process). Unfortunately, because of a change in staffing, the PSC has not been able to address the apparent discrepancies and uncertainties in the execution of the sampling, and it does not appear that it would be able to do so at this point.

The Council recognizes that problems of this kind can occur, particularly with the turnover of personnel, and when a lengthy period of time elapses between conducting the sampling and the final levels of administrative review. Given the strengths in the design of this audit in a number of respects (discussed below), and the fact that the methodology was fully documented and described by Dr. Moody in a written summary, the Council does not reach the decision to invalidate the extrapolation lightly. However, for the reasons explained above, errors and inaccuracies in executing the probability sample design, particularly in assigning claims to the strata and possibly in drawing the sample, make further errors and inaccuracies in the results of the sample highly likely. In this case, because of the nature and degree of the uncertainty about the sample, and the inability to correct the errors and inaccuracies at this point, the extrapolation is not reliable.

Valid Parts of The Probability Sample Process in This Case

As noted above, the Council does not agree with many of the bases the ALJ identified in his Analysis for invalidating the extrapolation. In fact, the Council agrees with a majority of the points made in the CMS memorandum about the validity of the sampling methodology used in this case. The Council identifies these points in the paragraphs that follow, because they are consistent with the provisions in CMS Ruling 86-1 and the relevant provisions in chapter 3 of the Medicare Program Integrity Manual. Thus, while the extrapolation cannot not be upheld because of the errors and inaccuracies identified above, the Council finds the following:

- The precision estimate of 15.67% in this audit is fully adequate to meet CMS requirements, as set forth in the Medicare Program Integrity Manual. There is no CMS or MPIM requirement for a 10% or lower precision estimate. See Exh. MAC-1 at 5-7. In fact, the MPIM's statistical sampling guidelines do not require any specific level of precision, but take into account all factors used in a particular statistical sampling methodology. Cf. MPIM, ch.3, §§ 3.10.4.1. (a number of different sampling designs are acceptable); 3.10.4.3. (sample size should be weighed together with other factors involved in sample design). In this case, a 15% precision percentage is reasonably low in comparison to those found in a number of cases the Council has reviewed. Moreover, the guidelines anticipate the assessment of an overpayment at the lower confidence bound of a one-sided 90% confidence interval. Cf. MPIM, ch. 3, § 3.10.5.1 (suggesting use of the lower limit of a one-sided 90% confidence interval as a conservative method). Use of the lower bound of a two-sided 90% confidence interval, as used here, is even more conservative and results in an even lower overpayment assessment.
- Additionally, the sampling error (or "coefficient of variation") in this case was computed at less than 10% in the calculations performed on both November 13, 2006, and on January 30, 2007. See Exh. MAC-3.⁵ This provides an

⁵ This eight-page document with additional statistical information on the sampling in this case, was filed at the back of the evidence file, in a section labeled "Non-Probative Correspondence Communication." The Council has made a copy of the document and placed it in the front of the file as Exh. MAC-3.

additional measure of precision in the sample, and further demonstrates that it is within an acceptable range.

- The ALJ appears to have confused or conflated the concepts of confidence interval and sampling precision. Dec. II at 2-3, 9; see also Exh. MAC-1 at 6-7. Sampling precision measures the degree of variability within the sample results in relation to the point estimate. The point estimate is the estimated total overpayment based on the single sample. MPIM, ch. 3, § 3.10.5.1, However, the confidence interval is determined by applying various t-values and/or z-values to obtain a desired (pre-selected) confidence interval, e.g. 90% one-sided, 90% two-sided, 95% two-sided, etc. Thus, a two-sided 90% confidence level, as calculated in this case, is not dependent on a particular precision level, because any desired confidence level can be calculated once certain data (e.g., the standard deviation and point estimate) are obtained from the sample results, and a t-value or z-value is selected based on the confidence level sought.
- The Council also disagrees with the ALJ's assessment of Dr. Haller's statements regarding distribution of the sample. The ALJ found that in employing the Central Limit Theorem, the existence of a skew in the sampled overpayment amounts in a single sample provides a reason to invalidate the single sample. See Dec. II at 3, 9; Exh. 3 at 4 (Dr. Haller's written testimony). As Mr. Landtroop testified, under the Central Limit Theorem, the issue would be whether *sample averages from multiple samples* are normally distributed, not whether a single sample has normal distribution. ALJ Hearing CD (Nov. 3, 2010) at 10:20 to 10:21 a.m. According to Mr. Landtroop, in the instant case, where only a single sample was taken and multiple sample averages are not available, there is no basis for invalidating a single sample based on skewness. *Id.*; see also Exh. MAC-1 at 7-8. The relevance of the Central Limit Theorem in this case, as in many of the overpayment cases involving statistical sampling, is that it demonstrates that a single sample of limited size (here, 60 claims) is sufficient to obtain a representative sample, even if the individual sample is skewed rather than normally distributed.
- In addition, insofar as the ALJ relied on Dr. Haller's logarithm adjustment to arrive at a different extrapolation

amount ("a correct extrapolation that was nearly one-third less," Dec. II at 9), the Council does not use that logarithm adjustment calculation as a basis for instead imposing a lower extrapolated overpayment recovery. Dr. Haller testified that he used natural logarithm transformation to render the overpayment data normally distributed. Exh. 3 at 4. While we note Dr. Haller's extremely strong credentials in statistics and do not question his knowledge in choosing this methodology, the Council has been unable to understand this logarithm process. Moreover, Mr. Landtroop raised some questions as to how certain adjustments for zero and negative numbers were made in this process. See ALJ Hearing CD (Nov. 3, 2010) at 10:21 to 10:22 a.m. (B.L. testimony); MAC-1 at 8. In any event, for the reason stated above, the skewed distribution of the single sample would not *per se* invalidate the sampling which was done here.

For all of the reasons stated above, the Council finds that the positions taken in the agency referral memorandum supporting the sampling and extrapolation process in this case were well taken. However, CMS's referral did not address the inaccuracies and uncertainties reflected in the data the PSC provided to the independent expert, which are documented in his written report. Thus, while the CMS memorandum points to methodology issues which the ALJ found to be material errors, the Council finds that many of these were not, in fact, material flaws in the chosen methodology in this case. However, due to errors and discrepancies in the conduct of the sampling and subsequent potential data errors, the Council invalidates the extrapolation which occurs in this case.

DECISION

For these reasons, the Council modifies the ALJ's decision. The Council concurs in the ALJ's decision to invalidate the extrapolation in this case, but for reasons that differ in some manner from the reasons given by the ALJ in his decision. The Council determines that the extrapolation is insufficiently reliable because of shortcomings in the way the samples were drawn or the frames were sorted, and concerns about the PSC's provision of inconsistent data to the independent expert reviewing the sampling without explanation to the statisticians

or to the Council. The appellant remains financially liable only for the overpayments on individual claims in the sample.

MEDICARE APPEALS COUNCIL

/s/ Gilde Morrisson
Administrative Appeals Judge

/s/ Constance B. Tobias, Chair
Departmental Appeals Board

Date: May 12, 2011