PROVIDER REIMBURSEMENT REVIEW BOARD DECISION

On the Record
2023-D33

PROVIDER–
Flowers Hospital

Provider No.:
01-0055

vs.

MEDICARE CONTRACTOR –
Palmetto GBA

RECORD HEARING DATE –
July 29, 2022

Fiscal Year Ending –
June 30, 2010

Case No. –
14-1468

INDEX

ISSUE STATEMENT

Whether the Medicare Contractor improperly calculated the Provider's Disproportionate Share Hospital ("DSH") reimbursement due to sampling errors in review of the Medicaid-eligible patient days.[1]

DECISION

After considering Medicare law and regulations, the arguments presented, and the evidence admitted, the Provider Reimbursement Review Board ("Board") finds that the Medicare Contractor's Medicaid-eligible patient days sampling methodology was valid, and that the Medicare Contractor properly calculated the Provider's DSH reimbursement for FY 2010.

INTRODUCTION

Flowers Hospital ("Flowers" or "Provider") is an acute care hospital located in Dothan, Alabama. Flowers' assigned Medicare contractor[2] is Palmetto GBA ("Medicare Contractor").

On August 20, 2013, the Medicare Contractor issued a Notice of Program Reimbursement ("NPR") for Flowers' fiscal year end ("FYE") June 30, 2010.[3] As part of its audit of Flowers' FY 2010 cost report, the Medicare Contractor reviewed a stratified sample from the total 7,197 Medicaid-eligible days reported by Flowers on its as-filed cost report.[4] The Medicare Contractor then extrapolated the error rate percentage resulting from its review, thereby disallowing 189 days from the total 7,197 Medicaid-eligible days. On August 20, 2013, the Medicare Contractor issued the FY 2010 NPR which reflected this disallowance and included an audit adjustment reducing Flowers' DSH payment by $98,801.[5]

Flowers timely appealed the Medicare Contractor's final determination and met the jurisdictional requirements for a Board hearing. On July 29, 2022, the Board approved Flowers March 25, 2022 request for a hearing on the record. Flowers was represented by Daniel Hettich, Esq. of King & Spalding, LLP. The Medicare Contractor was represented by Joseph Bauers, Esq. of Federal Specialized Services.

STATUTORY AND REGULATORY BACKGROUND: MEDICARE DSH PAYMENT

A. Medicare IPPS and DSH Reimbursement

Part A of the Medicare Act covers "inpatient hospital services." Since 1983, the Medicare program has paid most hospitals for the operating costs of inpatient hospital services under the inpatient

[1] Parties' Stipulation of Undisputed Facts (hereinafter "Stip.") at ¶ 1.4 (Mar. 25, 2022).
[2] CMS' payment and audit functions under the Medicare program were historically contracted to organizations known as fiscal intermediaries ("FIs"), but these functions are now contracted with organizations known as Medicare administrative contractors ("MACs"). The term "Medicare contractor" refers to both FIs and MACs as appropriate and relevant.
[3] Stip. at ¶ 1.5
[4] Stip. at ¶¶ 2.1, 2.2.
[5] Stip. at ¶¶ 2.3, 2.4.

prospective payment system ("IPPS").[6] Under IPPS, Medicare pays predetermined, standardized amounts per discharge, subject to certain payment adjustments.[7]

The IPPS statute contains several provisions that adjust reimbursement based on hospital-specific factors.[8] This case involves the hospital-specific DSH adjustment, which requires the Secretary to provide increased PPS payments to hospitals that serve a significantly disproportionate number of low-income patients.[9]

A hospital may qualify for a DSH adjustment based on its disproportionate patient percentage ("DPP").[10] As a proxy for utilization by low-income patients, the DPP determines a hospital's qualification as a DSH, and it also determines the amount of the DSH payment.[11] The DPP is defined as the sum of two fractions expressed as percentages.[12] Those two fractions are referred to as the "Medicare/SSI" fraction and the "Medicaid" fraction. Both fractions consider whether a patient was "entitled to benefits under part A."[13]

The Medicaid fraction is at issue and it is defined in 42 U.S.C. § 1395ww(d)(5)(F)(vi)(II) as:

> [T]he fraction (expressed as a percentage), the numerator of which is the number of the hospital's patient days for such period which consist of patients who (for such days) were eligible for medical assistance under a State plan approved under subchapter XIX [the Medicaid program], but who were *not entitled to benefits under part A of this subchapter,* and the denominator of which is the total number of the hospital's patient days for such period.[14]

The DSH regulation at 42 C.F.R. § 412.106(b)(4) (2009) specifies that Medicare contractors calculate the Medicaid fraction for a hospital's cost reporting period by "determin[ing] . . . the number of the hospital's patient days of service for which patients were eligible for Medicaid but not entitled to Medicare Part A, and divides that number by the total number of patient days in the same period."

B. Medicare Policy on Statistical Sampling

1. Usage of Statistical Sampling for Cost Report Audits

The Centers for Medicare and Medicaid Services ("CMS") published guidance for Medicare contractors on the use of statistical sampling for purposes of cost report audits in the Medicare Financial Management Manual, CMS Pub 100-06 ("MFMM") in Chapter 8 entitled "Contractor Procedures for Provider Audits."

---

[6] *See* 42 U.S.C. § 1395ww(d)(l)-(5); 42 C.F.R. Part 412.
[7] *Id.*
[8] *See, e.g.,* 42 U.S.C. § 1395ww(d)(5).
[9] *See* 42 U.S.C. § 1395ww(d)(5)(F)(i)(I); 42 C.F.R. § 412.106.
[10] *See* 42 U.S.C. §§ 1395ww(d)(5)(F)(v); 42 C.F.R. § 412.l06(c)(l).
[11] *See* 42 U.S.C. §§ 1395ww(d)(5)(F)(iv), (vii)-(xiii); 42 C.F.R. § 412.106(d).
[12] *See* 42 U.S.C. § 1395ww(d)(5)(F)(vi).
[13] *See, e.g.,* 42 C.F.R. § 412.106(b)(3), (4).
[14] (Emphasis added.)

Further, CMS has recognized that the statistical sampling software known as "RAT-STATS" is acceptable for Medicare contractors to use in assisting in developing and selecting random samples as well as extrapolating the sample results to the relevant universe being sampled.[15] RAT-STATS is a publicly available software package developed by the Office of the Inspector General of the U.S. Department of Health and Human Services ("HHS OIG").[16]

2. Usage of Statistical Sampling in Program Integrity Functions

On February 20, 1986, CMS, formerly known as the Health Care Financing Administration or "HCFA", issued HCFA Ruling No. 86-1 entitled "Use of Statistical Sampling to Project Overpayments to Medicare Providers and Suppliers" which is relevant to this appeal. This Ruling describes the Agency's policy for the use of statistical sampling to project overpayments to Medicare providers and suppliers in conducting Program Integrity activities. The Ruling states 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.[17]

Thus, the Ruling makes clear that a Medicare contractor's use of statistical sampling and the resulting extrapolated overpayment amounts have a presumption of validity and that a provider may challenge: (1) the statistical validity of the sample to rebut that presumption; and/or (2) the findings made as a result of the sample.

---

[15] For example, the MPIM § 3.10.4.2 (Rev. 282, Jan. 8, 2009) states: "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." *See also* Medicare Intermediary Manual, CMS Pub. 13-3, § 3940.3(B) (as revised by Transmittal 1770, Mar. 1, 1999) (stating "Identify the source of the random numbers used to select sample items. A recommended source of random numbers is RAT-STATS, although any reputable random number selection method may be used. RAT-STATS is a software application program that assists the user in selecting random samples and evaluating the results. The software is designed to operate on personal computers using Microsoft Disk Operating System (MS-DOS). The RAT-STATS software program was developed by the Department of Health and Human Services, Office of the Inspector General, Office of Audit Services, Regional Advanced Techniques.").

[16] The HHS OIG has posted information on RAT-STATS on its website at: https://oig.hhs.gov/compliance/rat-stats/ (last visited Aug. 31, 2023). The RAT-STATS information includes the 2010 User Guide and the 2010 Companion Manual.

[17] HCFA Ruling 86-1 at 11.

The Medicare Program Integrity Manual, CMS Pub 100-08 ("MPIM") includes guidance on statistical sampling both in Chapter 3 entitled "Verifying Potential Errors and Taking Corrective Actions" and in Chapter 8 entitled "Administrative Actions and Statistical Sampling for Overpayment Estimation." As part of the Medicare Modernization Act of 2003, Congress revised 42 U.S.C. § 1395ddd to include the following guidance at subsection (f)(3) on the use of statistical sampling for determining overpayments:

> (3) **Limitation on use of extrapolation**
>
> A Medicare contractor may not use extrapolation to determine overpayment amounts to be recovered by recoupment, offset, or otherwise unless the Secretary determines that –
>
> (A) there is a sustained or high level of payment error; or
>
> (B) documented educational intervention has failed to correct the payment error.
>
> There shall be no administrative or judicial review under section 1395ff of this title, section 1395oo of this title, or otherwise, of determinations by the Secretary of sustained or high levels of payment errors under this paragraph.[18]

As confirmed in the preamble to the final rule published on December 9, 2009, the Secretary implemented the above statutory provision by issuing a transmittal in June, 2005 to revise the MPIM instructions on the use of statistical sampling for overpayment extrapolation.[19]

---

[18] 42 U.S.C. § 1395ddd(f)(3) (2012).

[19] Specifically, the final rule entitled "Medicare Program: Changes to the Medicare Claim Appeal Procedures" published at 74 Fed. Reg. 65296, 65303-04 (Dec. 9, 2009) includes the following discussion regarding the Secretary's implementation of 42 U.S.C. § 1395ddd(f)(3):

> *Comment*: One commenter requested that we define the phrase "sustained or high levels of payment errors" (§ 405.926(p)) and requested that we specify how such determinations will be made. The commenter also requested that CMS review dismissals on the grounds that the claim involves a sustained or high error rate. The commenter suggested that CMS provide clarification of the implications of such a finding. Finally, the commenter recommended that CMS provide a mechanism for providers to be removed from this "sanction".
> *Response*: In section 1893(f)(3) of the Act, added by section 935 of the MMA, Congress placed restrictions on the use of extrapolation to determine overpayment amounts to be recovered from Medicare providers, suppliers or beneficiaries. In order to calculate an overpayment by extrapolation, there must be a determination of either: (1) A sustained or high level of payment error, or (2) a documented educational intervention that has failed to correct the payment error. In addition, in section 1874A(h)(2) of the Act, as added by section 934 of the MMA, Congress required contractors to identify a likelihood of sustained or high level of payment error under section 1893(f)(3)(A) of the Act before initiating non-random pre-payment reviews of a provider or supplier, and in section 1893(f)(3) of the Act, expressly precluded administrative or judicial review of contractor determinations of sustained or high levels of payment errors. Accordingly, we included a conforming provision at § 405.926(p) of the interim final rule providing that determinations of sustained or high levels of payment error are not initial determinations that may be appealed under this subpart. We note, however, that while the determination of whether a provider or supplier has a sustained or high level of

Specifically, MPIM Transmittal 114 revised MPIM, Ch. 3, § 10 entitled "Use of Statistical Sampling for Overpayment Estimation" for the purpose of "implement[ing] MMA Section 935(a), which amends [42 U.S.C. § 1395ddd] by adding new subsection (f)(3) Limitation on Use of Extrapolation."[20] Significantly, the Secretary did not otherwise revise Chapter 8 of the MFMM as part of the implementation of § 1395ddd(f)(3), presumably because it is not applicable to initial audits of as-filed cost reports.[21] As such, the discussions regarding statistical sampling in CMS Ruling 86-1, 42 U.S.C. § 1395ddd(f)(3) and the MPIM do not apply to this case. This conclusion is further supported by the fact that the statistical sampling at issue in this case was *not* used to extrapolate an overpayment. Rather, the statistical sampling was used to review and finalize the data to be used in one factor of a final payment rate calculation (*i.e.*, to review and finalize the number of Medicaid eligible days being used in the numerator of the Medicaid fraction which, in turn, is one factor used in finalizing the DSH adjustment calculation for FY 2010 in the NPR).

STATEMENT OF FACTS

This case focuses on the Medicaid fraction (percentage) as used in calculating DSH payments. Specifically, this case concerns the number of Medicaid eligible patient days to be used in the numerator of this fraction. These days are subject to audit and verification by the Medicare Contractor because, pursuant to 42 C.F.R. § 412.106(b)(4)(iii): "[t]he hospital has the burden of furnishing data adequate to prove eligibility for each Medicaid patient day claimed . . . , and of verifying with the State that a patient was eligible for Medicaid *during each claimed patient hospital day*."[22] This regulation codifies the following directive in HCFA Ruling 97-2, which , in relevant part, states:

> The hospitals bear the burden of proof and must verify with the State that a patient was eligible for Medicaid (for some covered services) *during each day* ***of the patient's inpatient hospital stay***. As the intermediaries may require, hospitals are responsible for

---

payment error is not subject to appeal, the initial or revised determinations made on the underlying claims for items or services would be subject to appeal.
CMS issued operating instructions for determining when a provider or supplier has a sustained or high level of payment error in June 2005: (http://www.cms.hhs.gov/transmittals/downloads/R114PI.pdf). Furthermore, we issued a final rule on September 26, 2008 (73 FR 55753) to address when contractors may terminate the non-random pre-payment review of claims submitted by a provider or supplier. The commenter's concerns regarding the practical considerations of determinations of a provider's or supplier's sustained or high error rates are beyond the scope of this regulation. With respect to the suggestion that CMS review dismissals on the grounds that the claim involved a sustained or high error rate, as noted above, while that determination does not constitute an initial determination and is not subject to appeal, any claim denials resulting from the review would constitute initial determinations that may be appealed. Therefore, we do not anticipate any denials of claims solely based on this determination. Rather, the determination of a sustained or high error rate will be used as the basis for a contractor undertaking further review of claims submitted by the provider or supplier. Finally, we strongly disagree with the commenter's characterization of the determination of a sustained or high error rate as a sanction. This determination does not result in an assessment of civil money penalties, or any other administrative action. Rather, it serves as the basis for a contractor's review of a provider's or supplier's subsequent claim submissions.

[20] MPIM Transmittal 114 at 1 (Change Request 3734, June 10, 2005).

[21] Indeed, there are no cross references to the MPIM in MFMM, Chapter 8.

[22] (Emphasis added.)

> and must furnish appropriate documentation to substantiate the number of patient days claimed. Days for patients that cannot be verified by State records to have fallen within a period wherein the patient was eligible for Medicaid cannot be counted.[23]

In its as-filed FY 2010 cost report, Flowers claimed 7,197 Medicaid-eligible patient days. In performing its audit of Flowers' claimed Medicaid-eligible patient days, the Medicare Contractor stratified Flowers' total Medicaid patient days into two stratums: an upper stratum of patient stays of 10 days or more (total population of 93 claims/1,916 days) and a lower stratum of patient stays of less than ten days (total population of 1,867 claims/5,281 days).[24]

For each stratum, the Medicare Contractor calculated the average and the standard deviation which were used to calculate the valid sample sizes using the RAT-STATS Software. The statistically valid samples sizes calculated to obtain 20 percent precision at a 90 percent confidence level were 10 claims from the upper stratum and 17 claims from the lower stratum. However, the Medicare Contractor's guidelines require a minimum sample size of 30; therefore, 30 claims were randomly selected for both the lower and upper stratums using the software's random number generator.[25]

The upper stratum sample was 30 claims (where each "claim" was a hospital inpatient stay) totaling 690 days in the aggregate. The Medicare Contractor found no errors in the sample for the upper stratum. Thus, all 1,916 days in the upper stratum population were found to be allowable. The lower stratum sample was 30 claims totaling 84 days in the aggregate. The Medicare Contractor found that a single patient with a three-day hospital stay was ineligible for Medicaid, resulting in an error rate of 3.57 percent in the sample taken from the lower stratum (*i.e.*, (3 days/84 total days in the sample) x 100). Applying an extrapolation method that extends the 3.57 percent error rate to the total lower stratum population of 5,281 days resulted in a disallowance of 189 days.[26]

The results of the Medicare Contractor's review are summarized in the following table[27]:

STRATUM	TOTAL STAYS & DAYS IN STRATUM	SAMPLE SIZE	SAMPLED DAYS ALLOWED	% SAMPLED DAYS DIS-ALLOWED	% SAMPLED DAYS ALLOWED	PROJECTED ALLOWABLE NO. OF DAYS
LOS ≥ 10 days	1,916 days	690 days across 30 stays	690 days	0%	100%	1,916 days
LOS < 10 days	5,281 days	84 days across 30 stays	81 days	3.57%	96.43%	5,092 days
Total						7,008 days

[23] HCFA Ruling 97-2 at 3 (Feb. 1997) (emphasis added.)
[24] Stip. at ¶ 2.2.1 and 2.2.2; Exhibit (hereinafter "Ex.") C-2 at 2, 3, 4.
[25] Ex. C-2 at 1.
[26] Stip. at ¶¶ 2.2.1, 2.2.2.
[27] Provider's Final Position Paper (hereinafter "Provider's FPP") at 9 (Sept. 2, 2021).

DISCUSSION, FINDINGS OF FACT, AND CONCLUSIONS OF LAW

Flowers argues that the "[Medicare Contractor]'s downward adjustment to the Provider's Medicare DSH payment based on extrapolation methods [is] contrary to both the plain language of the Medicare statute and sound statistical methods."[28] Flowers contends that "[t]he [Medicare Contractor]'s disallowance of 189 Medicaid patient days is improper and must be set aside for at least two principal reasons."[29] First, Flowers contends that "the [Medicare Contractor] extrapolated an error rate of only 3.57 percent that in no way constitutes 'a sustained or high level of payment error' that must be present [per 42 U.S.C. § 1395ddd(f)(3)] in order for the [Medicare Contractor] to extrapolate that error rate to a larger population."[30]

Flowers further asserts that the burden is on CMS, and its contractors, to establish these facts and that the Medicare Contractor's workpapers offer no support that the 3.57 percent error rate calculated during the audit "represents a sustained or high level of payment error, nor that educational intervention has failed to correct any payment errors."[31] Flowers continues its argument, stating:

> A reasonable person cannot conclude that an error rate of 3.57 percent – based on one ineligible patient out of 60 reviewed – meets the standards established by Congress in the [MMA] statute. For that reason, the [Medicare Contractor] may only disallow the three ineligible patient days it actually reviewed and should add 186 days back to the numerator of the Provider's Medicaid fraction.[32]

Next, Flowers asserts that the Medicare Contractor's disallowance was the result of "a flawed sampling methodology – namely, the decision to stratify the Provider's Medicaid patient population based on length of stay."[33] Further, Flowers argues that, "[b]y stratifying the claims based on length of stay, the [Medicare Contractor] has created strata with no inherent relationship to the variables being tested."[34] Specifically, Flowers described the stratification of its Medicaid patient population based on length of stay as illogical because it is "a characteristic that has no bearing on whether a patient is Medicaid eligible and entitled to benefits under Medicare Part A."[35] That improper stratification resulted in an error rate nearly ten times higher than what a simple sample of the 60 patient stays at issue would have revealed. Flowers claims its "true error rate is so insignificant as to obviate the need for an extrapolation. Had the [Medicare Contractor] not stratified its sample of the 60 stays under review, it would have arrived at an error rate of only 0.39 percent [3/774]."[36]

In support of its position, Flowers references a treatise on sampling techniques which states that three conditions "must be present for the effective and reliable use of stratified sampling. Stratification can produce "large gains in precision," but only if the following three conditions

---

[28] *Id.* at 2.
[29] *Id.* at 10.
[30] *Id.* at 3.
[31] *Id.* at 10-11.
[32] *Id.* at 12.
[33] *Id.* at 10.
[34] *Id.* at 13.
[35] *Id.* at 3.
[36] *Id.* at 12-13.

are satisfied: 1) the population is composed of institutions varying widely in size, 2) the principal variables to be measured are closely related to the sizes of the institutions, and 3) a good measure of size is available for setting up the strata."[37] In relation to these conditions, Flowers maintains:

> While the lengths of stay for each Medicaid patient did vary in size and were available for setting up the strata, the [Medicare Contractor's] methods lacked the second crucial element of Cochran's criteria. Specifically, the variables to be measured – whether the patients were eligible for Medicaid and not entitled to benefits under Part A – were not in any way related to the length of stay. A patient's eligibility for Medicaid is in no way related to or determined by the length of time that patient stayed in the hospital. Rather than improving precision, the [Medicare Contractor's] [sampling] methods produced an over-inflated error rate.[38]

Lastly, Flowers argues that the Board expressly rejected this type of sampling method in *Exempla Lutheran Med. Ctr.*, PRRB Dec. No. 2011-D32 (June 3, 2011) ("*Exempla*"),[39] wherein the Board "did not find any rational justification for the stratification of the sample based on length of stay" and "did not see a relationship between allowable DSH days and length of stay, and found that strata based on lengths of stay to be unnecessary."[40] Flowers asserts that the stratification at issue here is identical to that in *Exempla*. Therefore, "the Board should set aside the [Medicare Contractor's] stratification and the error rate derived from that stratification."[41]

The Medicare Contractor counters that the DSH payment adjustment, determined using sampling, "is justified as the findings are the direct result of the inaccurate population submitted by the Provider."[42] It argues that:

1. "Audit sampling is the application of audit procedures to less than 100 percent of the items within an account balance or class of transactions for the purpose of evaluating some characteristic of the balance or class" based on the premise that "[i]t is not cost effective to audit every transaction or account."

2. The use of sampling "is well accepted and addressed in numerous texts, publication, and manuals" and "allows for the stratification of the total universe into various populations based on the judgment of the reviewer."

3. It applied the statistical sampling method in the audit at issue "in accordance with the policies and provisions for sampling outlined in [MFMM, Chapter 8, § 60.6.B]."[43]

---

[37] *Id.* at 14.
[38] *Id.*; Ex. P-4.
[39] The CMS Administrator reviewed the Board's decision in *Exempla* and reversed the Board in his decision issued on July 27, 2011.
[40] Provider's FPP at 14 (referring to *Exempla* at 18).
[41] *Id.*; Ex. P-4.
[42] Medicare Contractor's Final Position Paper ("Medicare Contractor's FPP") at 7 (Sept. 30, 2021).
[43] *Id.* at 7-8.

The Medicare Contractor refutes Flowers' claim that extrapolation should not be used and asserts that Flowers has taken the directive in 42 U.S.C. § 1395ddd(f)(3) out of context. The Medicare Contractor asserts this provision relates to claim overpayments and recoveries and that "the sampling guidance provided in [the MPIM] should not be confused with the statistical sampling guidance included in [MFMM, Chapter 8].[44]

Chapter 8 of the MFMM provides guidance for Medicare Contractors in conducting audits of Medicare cost reports. Section 60.6 – Designing Tests/Sampling - provides guidance on sampling in the context of cost report audits. It reads, in pertinent part:

> Design such tests as are necessary to accomplish your audit objectives. Your tests must aid you in reaching conclusions necessary to complete the audit. Use sampling when this would be more efficient in testing the universe of transactions, entries, or statistical data within an area of consideration.
>
> Sampling is the application of an audit procedure to less than 100 percent of the items within an account balance, class of transactions, or statistics (e.g., count of interns/residents) to evaluate some characteristic of the such balance, class, or statistics. On the basis of facts known to you, decide if all transactions, balances, or statistics that pertain to the issue/area being tested need to be reviewed in order to obtain sufficient evidence. In most cases, an auditor will test at a level less than 100 percent.
>
> There are two general sampling approaches, nonstatistical and statistical. Either approach, when properly applied, can provide sufficient evidential data related to the design and size of an audit sample, among other factors. A nonstatistical sample may support acceptance of findings, but findings must be scientifically established to support adjustments.
>
> Some degree of uncertainty is inherent in applying audit procedures and is referred to as ultimate risk. Ultimate risk includes uncertainties due both to sampling and other factors. Sampling risk arises from the possibility that when a compliance or a substantive test is restricted to a sample, the auditor's conclusions may be different had the test been applied in the same way to all items in the account balance, class of transactions, or statistics.
>
> If you use a sample to test certain issues scoped for audit, you must include a description of the sampling technique, all parameters used to select the sample, and confidence level in the audit working papers.

[44] *Id.* at 10.

A. Planning Samples

Planning an audit involves a strategy for selecting appropriate sample(s). When planning a particular sample, consider:

- The relationship of the sample to the audit objective;
- Preliminary estimates of materiality levels;
- The allowable risk of incorrect acceptance; and
- Characteristics of the population, i.e., the items comprising the universe.

B. Selecting a Sampling Approach

Because either nonstatistical or statistical sampling can provide sufficient evidence, choose between them after considering their relative cost and effectiveness. Statistical sampling helps to:

- Design an efficient sample;
- Measure the sufficiency of the evidential matter obtained; and
- Evaluate the results.

By using statistical theory, quantify sampling risk in limiting to an acceptable level. Statistical sampling involves additional costs of designing individual samples to meet the statistical requirements and selecting items to be examined. Where the audit objective would be best accomplished by stratifying the universe/population into high and low strata (e.g., where Medicare bad debts are tested), use your judgment in designating the threshold for this stratification. Once determined, review all the items in the high strata population and use statistical or nonstatistical sampling to test the low strata.

C. Sampling Risk

In performing substantive tests of details, consider:

- The risk of incorrect acceptance that the sample supports the conclusion that the items are not materially misstated when they are; and
- The risk of incorrect rejection that the sample supports the conclusion that the items are materially misstated when they are not.

D. Using the Test Results

If the results of testing your sample that was selected using a nonstatistical method indicate probable errors in the universe of

> transactions, entries, or statistics, document your decision to expand the sample or redesign the sample using a statistical method. If the results of testing your sample that was selected using a statistical method indicate probable errors in the universe, document your decision to project the error to the universe/population.
>
> If your adjustment pertains only to the error(s) that was identified, you must document the reason for not considering the effect of the error(s) on the universe.[45]

As previously discussed, the Board agrees with the Medicare Contractor that neither 42 U.S.C. § 1395ddd(f)(3) nor Chapter 8 of the MPIM are applicable to this case as they apply to extrapolation of overpayments. The non-applicability of these provisions is further highlighted by the fact that the below MPIM excerpts are directed to "the PSC or ZPIC BI unit or the contractor MR unit."

However, the Board notes that the MPIM in effect at the sampling was conducted may provide some helpful guidance, and information, on *general* statistical sampling principles that transcend this specific context:

> **Section 8.4.2 – Probability Sampling**
>
> Regardless of the method of sample selection used, the PSC or ZPIC BI unit or the contractor MR unit shall follow a procedure that results in a probability sample. For a procedure to be classified as probability sampling the following two 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. . . .
>
> • Each sampling unit in each distinct possible sample must have a known probability of selection. . . .
>
> 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 sampling design is properly executed. i.e., defining the universe, the sampling 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 or that the resulting estimates are "not statistically valid" cannot legitimately be made. In other words, a probability sample and its results are always "valid." However,

[45] MFMM, Ch. 8, §§ 60.6, 60.6(B) (Rev. 60, Issued: Nov. 26, 2004) (copy at Ex. C-5).

because of differences in the choice of a design, the level of available resources, and the method of estimation, 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.

****

Section 8.4.4 – Sample Selection

Section 8.4.4.1 – Sample Design

Identify the sample design to be followed. The most common designs used are simple random sampling, systematic sampling, stratified sampling, and cluster sampling, or a combination of these.

****

8.4.4.1.3 Stratified Sampling

Stratified sampling involves classifying the sampling units in the frame into non-overlapping groups, or strata. The stratification scheme should try to ensure that a sampling unit from a particular stratum is more likely to be similar in overpayment amount to others in its stratum than to sampling units in other strata. . . . The independent random samples from the strata need not have the same selection rates. . . . The main object of stratification is to define the strata in a way that will reduce the margin of error in the estimate below that which would be attained by other sampling methods, as well as to obtain an unbiased estimate or an estimate with an acceptable bias. . . .

****

Section 8.4.4.4 – Documentation of Sampling Methodology

The PSC or ZPIC BI unit or the contractor MR unit shall maintain complete documentation of the sampling methodology that was followed.

****

Section 8.4.5 – Calculating the Estimated Overpayment

Section 8.4.5.1 – 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 provider . . . . 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 provider . . . . 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, the PSC or ZPIC BI unit or the contractor MR unit is not precluded from demanding the point estimate where high precision has been achieved.

****

Section 8.4.11 – Additional Discussion on Stratified Sampling and Cluster Sampling

Section 8.4.11.1 –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.

> . . . . 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. . . . 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. . . . *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*.[46]

The Board finds that the Medicare Contractor's sampling methodology is compliant with CMS instructions and guidance. The record shows that the Medicare Contractor used RAT-STATS (the HHS OIG statistical sampling software program discussed above) to identify sample sizes for each the 2 stratums, and then generated random sample selections from each respective stratum. Based upon the population in each stratum, the sample sizes were verified to be sufficient. The Medicare Contractor found an error in one stratum, then computed an error rate and extrapolated the error rate to the universe of that stratum.

As set forth in CMS guidance, the Medicare Contractor maintained complete documentation of the sampling methodology in its workpapers. The workpapers displayed the Medicare Contractor's sampling techniques, the confidence level of the sampling approach, and its decision to expand the error to the universe.

Based on its review of the Medicare Contractor's workpapers, the Board finds that it was not unreasonable for the Medicare Contractor to stratify the Medicaid-eligible days universe into an upper stratum of patient stays of 10 days or more and a lower stratum of patient stays of less than 10 days. The total universe consisted of 1,960 claims or patient stays. The upper stratum accounted for 36.73 percent of the universe and the lower stratum accounted for 63.27 percent of the universe. The Board notes that the average length of stay in the upper stratum was 20.60 days and the average length of stay in the lower stratum was 2.83 days. The average length of stay across the universe was 3.67 days.[47]

---

[46] MPIM, Ch. 8, §§ 8.4.2-8.4.11 (Trans. 377, Rev. May 27, 2011) (italics emphasis added). Note the excerpts contain not only *general* statistical sampling principles but also procedural or process instructions that may be only applicable to overpayments.

[47] Ex. C-2.

The Medicare Contractor's workpapers demonstrate how the Medicare Contractor used the RAT-STATS variable appraisal module to determine the sample size sufficient to obtain 20 percent precision at a 90 percent confidence level. However, the Medicare Contractor determined that the resulting sample size of each stratum did not meet their organizational sampling policies (even though they did meet industry standards) and went further, resulting in the random examination of 30 patient stays per strata. Flowers makes no allegation that the sample sizes were too small.

While the use of length of stay to determine the upper and lower strata *may* not have had an optimal relationship with the survey variable,[48] the Board finds that this approach does not, on its face, mean that the strata and samples taken therefrom were biased. MPIM, Ch. 8, § 8.4.11.1[49] confirms the general sampling principle that stratification itself does not introduce bias and its use will generally ***reduce*** sampling error:

> 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 circumstances proportionate allocation ***will reduce*** *the sampling error* over that for an unstratified simple random sample.[50]

The fact that the Medicare Contractor did not explain why it decided to end the lower stratum at 9 days and begin the upper stratum at 10 days does not raise any concerns for the Board, much less create any fatal flaws. As noted above, the Medicare Contractor's use of the 2 strata presumably *reduced* the sampling error in comparison to an unstratified random sample and, thus, the use of the strata did not introduce bias.

Moreover, Flowers apparently fails to understand that a patient's Medicaid eligibility can change during the course of an inpatient hospital stay (usually where the patient stay straddles one or

---

[48] For example, in analyzing the distribution of lengths of stay across the universe of claims, maybe it would have been more optimal to use 3 strata (such as one for lengths of stay 5 days of less, another for lengths of stay 6 to 14 days, and another for lengths of stay greater than 14 days).

[49] While the MPIM is not generally applicable to cost report audits, the general sampling principle stated here is applicable as it is not specific to the program integrity context.

[50] (Emphasis added.) The following excerpt from Exhibit P-5 (Practical Statistical Sampling for Auditors by Arthur J. Wilburn) at 6 essentially states the same idea and notes that stratification generally increases precision:

> If the auditor knows nothing about the universe structure except its approximate size, it is best to select a simple random or systematic sample. It is very rare, however, that an auditor knows nothing about an accounting universe. He knows from his survey that the universe consists of different kinds of transactions and values which are likely to show marked differences in characteristics. Thus, he should endeavor to use this knowledge and other information about the universe to improve the efficiency of the sample design.
>
> Often, supplementary information about the universe permits the auditor to separate it into a number of groups or layers, and to select a random sample from each group, called a stratum. This is called stratified random sampling. *This procedure usually increases the precision of the sample result*. The degree of improvement, however, is dependent upon the skewness of the distribution of the universe and how skillfully the stratification is performed.
>
> Stratification tends to isolate or separate the more extreme possibilities which may occur under simple random sampling.

(Emphasis added.)

more months since Medicaid eligibility often begins and/or ends on a monthly basis). Further, as explained at HCFA Ruling 97-2 (as codified at 42 C.F.R. § 412.106(b)(4)(iii)), "The hospitals bear the burden of proof and must verify with the State that a patient was eligible for Medicaid (for some covered services) *during each day* ***of the patient's inpatient hospital stay***."[51]

Here, the Medicare Contractor audited the Medicaid-eligible days, stratified by the length of the patients' hospital stay, which is consistent with HCFA Ruling 97-2.[52] The lower stratum included lengths of stay that ranged from 1 day to 9 days, and the upper stratum included lengths of stay that ranged from 10 days to 118 days. It is intuitive that the nature of the Medicare Contractor's review is different with longer lengths of stay versus shorter lengths of day.[53] Further, an error in an outlier (*e.g.*, invalidating some or all of a very long length of stay such as the 118-day length of stay) could potentially have a disproportionate impact when projected on the universe. The Medicare Contractor's use of the two strata clearly reduced that potential as it separated the more extreme possibilities which may have occurred if simple random sampling had been used.

In this regard, the Board rejects Flowers' argument that the Board decision in *Exempla* provides any valid guidance or *persuasive* authority.[54] The facts in that case are distinguishable from the instant case and, further, Flowers misconstrues the Board's discussion on stratification.[55] The Board recognizes that, in *Exempla*, it "d[id] not find any rational justification for all the stratifications of the sample" and found that "the breakdown of two strata based on length of stays less than ten days and stays 10-47 days unnecessary."[56] However, the Board did not find the use of *stratification alone* to be a fatal flaw. Rather, as noted above, stratification generally does reduce error. Moreover, *Exempla* focused on the provider's argument that the Medicare Contractor's disallowance of the DSH payment for the fiscal period in dispute should be reversed because the Medicare Contractor's disallowance was derived from a "judgmental sample," which is not a statistically valid random sample. That is not the case here. Outside of its opposition to the stratification, Flowers makes no argument that the samples were not statistically valid random samples. Indeed, the Board notes that, to obtain 20 percent precision at a 90 percent confidence level, a sample of only 17 claims was needed for the lower stratum and the Medicare Contractor increased the randomly selected sample to 30, thereby demonstrably increasing the precision. In contrast, if the Board were to eliminate the stratification and project the result from the two samples onto the universe (as Flowers advocates), the Board would be introducing bias, decreasing precision, and causing the samples to no longer be statistically valid since the samples were *not*

---

[51] (Emphasis added.)
[52] *See infra* notes 53, 59 and accompanying text.
[53] It is intuitive to audit by hospital stay since all the days associated with a hospital stay are likely to have the same finding on Medicaid eligibility. In other words, if the patient had Medicaid eligibility during the hospital stay, he/she likely had it for the whole stay. That said, the longer the hospital stay the greater the likelihood that the patient's Medicaid eligibility status changed during the course of the hospital stay (whether becoming eligible or losing eligibility). This would support the use of strata to differentiate between shorter stays and outliers or longer stays. Thus, it makes sense to use the hospital stay as the sampling unit and to stratify based on length of stay.
[54] Neither Board nor Administrator decisions have mandatory precedence. *See* Provider Reimbursement Manual, CMS Pub. 15-1, §§ 2916, 2927(C)(6)(e).
[55] Moreover, it should not be lost that the Administrator reversed the Board as discussed at *supra* note 39 and accompanying text.
[56] PRRB Decision 2011-D32 at 17 (June 30, 2011).

randomly selected from the ***full*** universe and would, thereby, no longer have the same probability of being randomly selected.[57]

In support of its position, Flowers has included, in the record, certain excerpts from two treatises on sampling techniques. However, these excerpts are choice one-page excerpts without context and, thus, provide little, or no, evidentiary value.[58] Further, Flowers chose not to provide any expert testimony, such as from a statistical expert, to demonstrate why the Medicare Contractor's sampling, stratification, and extrapolation were objectively invalid. The Board finds that Flowers did not meet either its burden of production of evidence or its burden of proof in this regard.[59] Further, Flowers does not dispute the Medicare Contractor's findings based on the sample from the lower stratum.

The Medicare Contractor conducted its audit in accordance with the directives in MMFM, Chapter 8. Those directives require that the Medicare Contractor plan and perform the audit to obtain reasonable assurance that the cost report settlement reflects payment amounts and financial data in accordance with Medicare laws, regulations, and instructions.

As the MPIM emphasizes, 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. Again, the fact that the Medicare Contractor utilized a sampling methodology that another statistician may not prefer does not provide a basis for invalidating the sampling or the extrapolation as drawn and conducted.

Accordingly, the Board finds that Medicare Contractor's stratification method is not a fatal flaw and, thus, does not invalidate the sampling. To hold otherwise would be to ignore the reality of constraints imposed by conflicting demands on limited government funds, constraints which CMS chose to incorporate into the statistical sampling guidelines.

DECISION

After considering Medicare law and regulations, the arguments presented, and the evidence admitted, the Board finds that the Medicare Contractor's Medicaid-eligible patient days sampling methodology was valid, and that the Medicare Contractor properly calculated Flowers' DSH reimbursement for FY 2010.

---

[57] *See, e.g.,* MPIM, Ch. 8, §§ 8.4.2, 8.4.4.1.3 (discussing probability sampling and that probability of selection may properly vary by strata).

[58] Exhibit P-4 is a one-page excerpt (page 101) from "Statistics: Sampling Techniques" by William Cochran (3rd Ed. 1977) and Exhibit P-5 is 4 different one-page excerpts (pages 57, 66, 100, and 108) from "Practical Statistical Sampling for Auditors" by Arthur J. Wilburn (1984).

[59] *See* 42 C.F.R. § 405.1871(a)(3) (stating, in pertinent part: "The [Board hearing] decision must include findings of fact and conclusions of law regarding . . . whether the provider carried its burden of production of evidence and burden of proof by establishing, by a preponderance of the evidence, that the provider is entitled to relief on the merits of the matter at issue.

Board Members Participating:

Clayton J. Nix, Esq.
Robert A. Evarts, Esq.
Kevin D. Smith, CPA
Ratina Kelly, CPA

For the Board:

9/20/2023

X Clayton J. Nix

Clayton J. Nix, Esq.
Chair
Signed by: PIV