REPORT ON

ASSESSMENT OF AUDIT SAMPLING AND EXTRAPOLATION PROCESS

WASHINGTON HEALTH CARE AUTHORITY
DIVISION OF PROGRAM & PAYMENT INTEGRITY
OFFICE OF PROGRAM INTEGRITY

Lewis and Ellis, Inc. – Actuaries & Consultants
David Dillon, FSA, MAAA, MS
Ari Loiben, MS

Healthcare Actuaries
Roger Burton, FSA, MAAA, FCA

JANUARY 2014
Assessment of Audit Sampling and Extrapolation Process
January 2014

Executive Summary

The Office of Program Integrity (OPI), within the Washington Health Care Authority (HCA), makes payments to providers of medical goods and services provided to Medicaid qualified recipients. As part of the payment process, OPI conducts post-payment reviews to validate payments made.

As part of the post-payment review, OPI may conduct an on-site claims audit to determine the correct amount that should have been paid for services provided. Due to the typically large number of claims paid to a single provider, an audit of the entire population of claims is not administratively or financially viable. As a result, OPI uses statistical sampling and extrapolation methods to determine a subset of provider claim payments to audit. Based on this procedure, OPI estimates the amount of claim payments to be recovered from the provider.

Lewis & Ellis, Inc. (L&E) was engaged by OPI to assess the appropriateness and validity of the “combined stratified regression interval estimator”, which is the primary statistical approach utilized by OPI. Pursuant to WAC 182-502A-0900, OPI may employ other statistical methods in certain situations. Those methods were not evaluated.
Executive Summary

Sampling

When there is a large volume of claims involved in a population, auditing the entire population is typically not viable. Therefore, pursuant to WAC 182-502A-0900, a random sample of claims can be selected for the audit.

OPI utilizes a statistical method known as stratified random sampling and selects a total sample size based on a traditional statistical formula (subject to a minimum amount selected). The ultimate sample size depends on the number of claims within certain dollar amount ranges, the variation of claim amounts, and the levels of precision and confidence desired.

OPI typically uses a sample size formula that produces with 95% confidence that the upper bound of the estimated claim amount is within 5% of the estimated total correct claim amount.

Extrapolation

Once the sample size and sampling approach is determined, the claims to be sampled for the audit are chosen. The results of the claim audits are then extrapolated by using a regression estimator and a Normal approximation with a 95% confidence limit. The approach creates a “fair” estimate in which the risk of underpayment to a provider is 5%.

Summary

L&E concludes that the “combined stratified regression interval estimator” utilized by OPI is appropriate, reasonable, correctly applied, and meets industry standards for statistical auditing. L&E’s conclusion is based on a review of published research as well as independent validation of the methods employed by OPI using data provided by OPI for this assessment. It should be noted that if the information and data received from OPI is flawed, then the validation process and conclusions may not be valid.
## Contents

Chapter 1 Introduction ........................................................................................................ 1-1  
  PURPOSE AND SCOPE ..................................................................................................... 1-1  
  AUDIT AUTHORITY ......................................................................................................... 1-2  
  RELIANCE AND LIMITATIONS OF THIS STUDY .......................................................... 1-3  

Chapter 2 Sampling of Claim Payments ........................................................................... 2-1  
  INTRODUCTION ................................................................................................................ 2-1  
  OPI SAMPLE METHOD ..................................................................................................... 2-4  
  L&E ASSESSMENT OF SAMPLE METHOD ....................................................................... 2-11  

Chapter 3 Extrapolation of Audit Results ....................................................................... 3-1  
  INTRODUCTION ................................................................................................................ 3-1  
  OPI EXTRAPOLATION METHOD ...................................................................................... 3-3  
  L&E ASSESSMENT OF EXTRAPOLATION METHOD ...................................................... 3-5  

Chapter 4 Conclusion ......................................................................................................... 4-1  
  GENERAL ASSESSMENT ................................................................................................. 4-1  
  CONSISTENCY WITH STATUTORY AUTHORITY ............................................................ 4-2  

Appendix A ASOP 41 Disclosures ..................................................................................... 4-1
Chapter 1
Introduction

PURPOSE AND SCOPE

The Office of Program Integrity (OPI), within the Washington Health Care Authority (HCA), makes payments to providers of medical goods and services provided to Medicaid qualified recipients. In order to provide a “fair” estimate of the amount that should have been paid to a provider by OPI, a post-payment claim audit may be conducted. As part of the payment process, OPI conducts post-payment reviews to validate payments made.

As part of the post-payment review, OPI may conduct an on-site claims audit to determine the correct amount that should have been paid for services provided. Due to the typically large number of claims paid to a single provider, an audit of the entire population of claims is not administratively or financially viable. As a result, OPI uses statistical sampling and extrapolation methods to determine a subset of provider claim payments to audit. Based on this procedure, OPI estimates the amount of claim payments to be recovered from the provider.

This report assesses the appropriateness and validity of the statistical procedures utilized by OPI. Specifically, L&E was asked to complete an objective assessment of OPI’s statistical program and process by evaluation and analysis to include but not limited to the validation of audit methodologies including sampling and extrapolation.
AUDIT AUTHORITY

OPI is responsible for maintaining an ongoing program to audit providers participating in HCA programs. This responsibility and authority is governed by multiple Federal and State laws and regulations.

Federal Authority

The Federal rules governing proper payments and other program integrity activities include:

- Federal Executive Order No. 13520 (11/20/09);
- Improper Payments Information Act of 2002;
- Presidential Memorandum Regarding Finding and Recapturing Improper Payments (3/10/2010);
- 42 CFR 431.10 State agency designee to administer Medicaid program;
- 42 CFR 447.202 Payment for services – Audits;
- 42 CFR 455 Requirements for State fraud detection;
- 42 CFR 438 Managed Care;
- 42 CFR 465 Utilization control; and
- 42 CFR 1001 Program Integrity.

State Authority

The State laws and regulations governing program integrity include:

- RCW 74.09.200 Audits and investigations – Legislative declaration – State authority;
- RCW 74.09.290 Audits and investigations of providers – Patient records – Penalties;
- WAC 182-502A Provider audits and appeals; and
- WAC 182-502-0230 Provider payment reviews and dispute rights.

Specifically regarding audit sampling, extrapolation, and claim review, WAC 182-502A-0900 states:
(1) The department's procedures for auditing providers may include, but are not limited to, the following:
   (a) The use of random sampling and extrapolation; and/or
   (b) A claim-by-claim based review.
(2) The department's sample sizes are sufficient to ensure a minimum of ninety-five percent confidence level.
   (a) When calculating the amount to be recovered, the department totals all overpayments and underpayments reflected in the sample and may extrapolate to the universe from which the sample was drawn.
   (b) When the department uses the results of an audit sample to extrapolate the amount to be recovered, the provider may request a description of all of the following:
      (i) The universe from which the department drew the sample;
      (ii) The sample size and method that the department used to select the sample; and
      (iii) The formulas and calculation procedures the department used to determine the amount of the overpayment.
   (c) If a provider rebills a claim(s) for an adjustment and that claim(s) is part of the audit universe, the department does not remove the original paid claim(s) amount from the audit universe.
(3) When a claim-by-claim audit is conducted, specific claims are selected from the universe and audit overpayments are not extrapolated.
(4) The department recovers overpayments identified in the final audit report.
(5) The department does not consider non-billed or zero paid services or supplies when calculating underpayments or overpayments.
(6) The department considers undocumented services to be program overpayments.

RELIANCE AND LIMITATIONS OF THIS STUDY

This report has been prepared for the use of the Office of Program Integrity, within the Washington Health Care Authority to assess its current procedures regarding audit sampling and overpayment extrapolation. It is not appropriate for any other purpose.
In performing this assessment, L&E relied on data and information from many sources. L&E did not audit the data sources for accuracy, although the data was reviewed for reasonableness. If the data or information provided was inaccurate, then the assessments provided may be erroneous.

The primary author of this report is a member of the American Academy of Actuaries and meets the qualification standards for conducting actuarial services concerning healthcare issues. Both the primary and secondary authors of this report have received Masters of Statistical and/or Actuarial Sciences. The guidance expressed in this report are those of the authors only and do not necessarily represent the opinions of other L&E consultants within the firm.

The authors of this report are not attorneys and they are not qualified to give legal advice. Users of this report should consult legal counsel for interpreting legislation and administrative rules, and other issues related to the claims audit process.

The authors of this report are aware that it may be distributed to third parties; however, any users of this report must possess a certain level of expertise in health care, statistics, and/or actuarial science so as not to misinterpret the analysis and assessments. Any distribution of this report must be made in its entirety. In addition, any third party with access to this report acknowledges, as a condition of receipt, that L&E makes no representations or warranties as to the accuracy or completeness of the material. Any third party with access to these materials cannot bring suit, claim, or action against L&E, under any theory of law, related in any way to this material.
Chapter 2
Sampling of Claim Payments

INTRODUCTION

When there is a large volume of claims involved in a population, auditing the entire population is typically not viable. Therefore, pursuant to WAC 182-502A-0900, a random sample of claims can be selected for the audit.

Sampling: Stratified and Non-Stratified

A non-stratified random sample (a.k.a. simple random sample) is a straightforward technique that ensures that the sample includes representation from the entire population. That is, each eligible claim has an equally likely chance of being selected to be audited.

When taking a simple random sample of claims, the selection process is performed using a random number generator, which in turn, is used to select the sample from the entire population for a given provider. The size of the sample will depend on the desired level of precision and the assumed level of variation in the dollar values of the population of claims.

An alternative technique to simple random sampling is stratified random sampling. In this technique, the population is separated into smaller subsets called strata. Once the subsets are established, simple random samples are chosen from each subset.

A reason to use stratified random sampling instead of simple random sampling is that:

- The overall sample size can be reduced while maintaining the same level of precision; or
- The precision can be increased using the same overall sample size.

Additionally, when stratification is employed, the sizes of the random samples can vary from stratum to stratum. One approach is to sample from a stratum according to the size of the stratum (known as, proportional allocation). A second approach is to sample from a stratum according to both the size of
the stratum and the level of variation in claim amounts within the stratum (known as, optimum allocation).

The optimum allocation technique can provide for higher precision than the proportional allocation technique. Optimal allocation is the method used by OPI.

Creating the Strata

The key considerations and questions in creating the strata include:

- What is the best characteristic for the construction of the strata;
- How should the boundaries be determined;
- How many strata should there be; and
- How homogeneous is the data.

Stratification Variables

There are several approaches in determining which variable to use in order to create strata. These could include, but are not limited to, the book amount of claims, the medical procedure codes for the service provided, and the date of service.

Sample Size

The sample size is influenced by a number of factors, including the purpose of the study, the population size, the risk of selecting a "bad" sample (e.g., a sample not representative of the population), and the allowable sampling error.

That is, the criteria to use in determining the appropriate sample size include:

1. The level of precision;
2. The level of confidence or risk; and
3. The degree of variability in the attributes being measured.
**The Level of Precision**

An estimate’s level of precision is a measure of how close the estimate is to the quantity being estimated. The difference in the estimate and the quantity being estimated is referred to as the sampling error. One way to express sampling error is by using the margin of error. The margin of error is a measure of the precision of a sample estimate of the population value and is defined as the range in which the true value of the population is estimated to fall.

This range is often expressed in percentage points, (e.g., ±5 %). For example, if it is determined that 75% of the sample has a certain attribute, it would be concluded that between 70% and 80% of the entire population would have that same attribute.

**The Level of Confidence or Risk**

The confidence or risk level is based on ideas encompassed under the Central Limit Theorem. The key idea is that when a population is repeatedly sampled, the average value of the attribute obtained by those samples is equal to the true population value. Furthermore, the values obtained by these samples are distributed Normally about the true value.

In a Normal distribution, approximately 95% of the sample values are within two standard deviations of the true population value (e.g., a mean or median). This implies that if a 95% confidence level is selected, 95 out of 100 samples will have the true population value within the range of precision specified.

**The Degree of Variability**

The degree of variability measures the dispersion of the actual values in the dataset. The larger the variability in a population, the larger the sample size required to obtain a given level of precision. When a dataset has limited variation, a smaller sample size is acceptable.

For example, assume that the average value of a set of claims is to be estimated. If the range of values is between $100 and $125, a reasonable estimate of the average amount can be approximated after a few samples. However, if the range of claim amounts is $100 to $1,000,000, a much larger sample will be required to calculate a precise estimate of the average.
OPI SAMPLE METHOD

OPI uses a stratified sampling technique and selects a total sample size based on a traditional statistical formula subject to a minimum amount selected. The ultimate overall sample size value of \( n \) depends on the number of claims in each stratum, the variation of claim amounts, and the levels of precision and confidence desired.

The OPI typically uses a sample size formula that will produce with 95% confidence that the upper bound is within 5% of the estimated total correct claim amount.

Sampling Units

Sampling units are the elements selected for sampling. When auditing providers, the sampling units could include, but are not limited to, claims, encounters, beneficiaries, or claim lines.

Sampling units for OPI’s audits are claim lines. Although the terms “claims” or “claim amounts” may be used in the discussion to follow, the terms are considered interchangeable with claim line.

For most providers, the claim line is uniquely determined by both the claim number (TCN) and the line item number (line number). Exceptions to this would be for some providers, such as pharmacists, where there is only one line per claim. In this exception, the claim number uniquely determines the claim line.

The sampling frame is a listing of all possible sampling units from which the sample is selected. Only claim lines eligible for sampling are considered. Eligibility criteria require non-zero payment, non-dual beneficiary, specified provider, and the specified time period.

In addition to the TCN and line number, the recipient field is occasionally used. The TCN is a numeric field consisting of 15 or 18 digits, the line number is a 3 digit integer field, and the recipient ID is an 11 digit alpha-numeric field. Finally, the “reimbursement amount at line” (or reported amount) is the amount that was paid to the provider by OPI for that specific claim line.
Creating the Strata

The method used by OPI was developed to produce the “best” strata boundaries. “Best” is defined as: for a given number of strata and the overall sample size, the strata boundaries will produce the most precise estimate of the total correct audited claim amounts.

Pursuant to WAC 182-502A-0900, the department's procedures for auditing providers may include a claim-by-claim based review. When a claim-by-claim audit is conducted, those specific claims are selected from the universe and audit overpayments are not extrapolated. OPI uses a combination of stratified random sampling and claim-by-claim audits.

For the statistical process used by OPI, the universe of claims by provider is sorted by claim amount. OPI separates out the twenty-five claims with the largest reimbursement amounts. This set of claim lines will be referred to as the “Top 25” or stratum number 21. Each claim in this stratum is automatically designated for an audit, and the results are not included in the statistical extrapolation.

To identify the claim lines in stratum 21, the claims are sorted by reported amount, then by descending TCN, and then by descending line number. Therefore, if more than one claim line has an amount equal to the 25th greatest amount, then the claim lines associated with the least TCN value will be selected for the Top 25. If there is more than one claim line with the same TCN with an amount equal to the 25th greatest amount, then the claim line with the least line number will be selected for stratum 21.

The remaining claims are then placed into twenty “major strata.” The objective is to create strata whereby the reported amounts within each stratum are relatively homogeneous in nature.

Establishing the strata boundaries for the other 20 strata is done iteratively. The claims lines are re-sorted by reported amount, ascending TCN, and then ascending line number.

Trial Strata

The twenty-sixth largest reported amount and the smallest reported amount are defined by $M$ and $m$, respectively. The variable $d$ is defined to represent their difference (i.e., $d = M - m$). One
hundred trial strata are obtained by dividing this difference into 100 intervals of equal dollar width. Then, claim lines with amounts less than \( \frac{1}{100}d + m \) are included in the first trial stratum. Claim lines with amounts greater than or equal to \( \frac{1}{100}d + m \) but less than \( \frac{2}{100}d + m \) are included in the second trial stratum. This process is continued iteratively. Note that this process could produce trial strata that are empty.

**Detail Strata**

Once the trial strata are created, detail strata are created by subdividing any trial stratum that contains more than 3% of the claims. The trial stratum from which a detail stratum is constructed is called the “parent” of that detail stratum. The subdivision takes place by starting with the claim lines having the lowest dollar value within that particular stratum.

If the subdivision process proceeds without a detail stratum containing more than 3% of all claims, then the last detail stratum, within a given trial stratum, will typically contain whatever is left after the last full 3% detail stratum is constructed. Therefore, the last detail stratum constructed from a trial stratum will contain fewer than 3% of all claim lines.

**Major Strata**

The next step in the process is to put the detail strata in ascending order by reported amount of the claim lines within each stratum. That is, the stratum containing claims of the smallest dollar amounts is numbered one. Variable \( f_i \) is defined as the number of claims in detail stratum \( i \). The detail strata are produced in one of two ways:

1. A detail stratum either corresponds directly to a trial stratum that was not subdivided (that is, it did not contain more than 3% of all claims), or
2. It is the result of its parent trial stratum being subdivided.

If the detail stratum \( i \) corresponds directly to a trial stratum that was not subdivided, then \( u_i = 1 \). If the parent stratum for detail stratum \( i \) was subdivided into \( t \) parts, then \( u_i = \frac{1}{t} \). The formula \( \sqrt{u_1f_1} + \sqrt{u_2f_2} + \cdots + \sqrt{u_if_i} \) is defined as the “cumulative square root of frequency for detail stratum \( i \)”. Variable \( F \) is defined as the cumulative square root of the frequency for the
highest numbered detail stratum (i.e., \( F = \sum_{all\ strata} \sqrt{u_if_i} \)). To create the first major stratum, the detail stratum that has a cumulative square root of frequency that is closest to \( \frac{1}{20}F \) is selected. This stratum and all detail strata lower than this stratum will be included in the first major stratum.

In general, for the \( i^{th} \) major stratum, the detail stratum that has a cumulative square root of frequency that is closest to \( \frac{i}{20}F \) is selected. This stratum and all detail strata lower than this stratum that have not already been allocated to a major stratum will be included in the \( i^{th} \) major stratum.

**Total Sample Size**

Once the claims have been allocated to the twenty strata from which random samples are to be drawn, the sample size \( n_i \) of each of these random samples must be determined. The method involves allocating a certain proportion of the entire sample to each stratum. Therefore, it is necessary to first determine the total sample size \( n \).

The total sample size \( n \) is determined by a standard formula subject to a preset minimum value of 300. The formula is defined as follows: \( n = \frac{(1.1)(\sum N_iS_{xi})^2}{X^2(P^2/\varepsilon) + \sum N_iS_{xi}^2} \), where the summations are taken over all strata. \( X \) represents the sum of reported amounts for the entire population, \( N_i \) the population size of stratum \( i \), \( P \) the desired level of precision, \( C \) a factor associated with the desired confidence level, and \( S_{xi}^2 \) the variance of the reported amounts in stratum \( i \).

This formula for total sample size is identical to that presented by Cochran\(^1\) for optimum allocation with the exception of the 1.1 “safety factor”. The default values for \( P \) and \( C \) are 0.05 and 1.6449 (e.g., 95% confidence level), respectively. These default values frequently produce a calculated sample size which is overridden by the minimum sample size of 300.

Strata Sample Sizes

The total number of claims in stratum $i$ is denoted by $N_i$. Let $x_{ij}$ represent the book amount of claim $j$ from stratum $i$. Some of the formulas to follow will call for adding up the claim amounts from the sample, while others will call for the sum of all claim amounts. That is, claim $j$ from stratum $i$ may represent the $j^{th}$ claim sampled from stratum $i$, or the $j^{th}$ claim from among all claims in stratum $i$. A note following each formula will clarify the situation.

Similarly, $y_{ij}$ will represent the correct amount of claim $j$ from stratum $i$. Note that if a claim is audited, the correct amount is also called the audited amount. If $y_{ij}$ represents the correct amount of a claim that is not audited, then this quantity is never known. In that case, the quantity $y_{ij}$ is of theoretical rather than computational value.

The “true mean” and the “true variance” for the claims in stratum $i$ are given by $\bar{Y}_i = \frac{\sum y_{ij}}{N_i}$ and $S^2_{yi} = \frac{\sum (y_{ij} - \bar{Y}_i)^2}{N_i - 1}$, respectively, where both summations are over all claims in stratum $i$. Since $S^2_{yi}$ is unknown, its value is estimated using the quantity $S^2_{xi} = \frac{\sum (x_{ij} - \bar{X}_i)^2}{N_i - 1}$, where the summation is taken over all claims in stratum $i$, and $\bar{X}_i$ is the average reported value of all claims in stratum $i$.

The choice of sample size for each stratum is based on the optimum allocation of the total sample size given by the Neyman² allocation. The variable $n_i$ is defined as $n_i = n \frac{N_i S_{yi}}{\sum N_k S_{yk}}$ where the summation is taken over all strata and $n$ is the total number of claims to be sampled from all strata. Again, the true variance for stratum $i$ must be estimated using the variance of the reported values for stratum $i$, i.e., $S^2_{xi}$. Therefore, the value $n_i = n \frac{N_i S_{xi}}{\sum N_k S_{xk}}$ is used, where the summation is taken over all strata.

The final step used by OPI is to make sure the sample size is at least 8 and is defined by $n_i = \min(N_i, \max(8, n \frac{N_i S_{xi}}{\sum N_k S_{xk}}))$. This formula indicates that the sample size can never exceed the

---

actual number of claims in the stratum $N_i$ and the sample size will be at least 8 (if the stratum has at least that amount). Requiring the minimum sample size to be 8 is intended to reduce the impact that one or two extreme claims may have on the estimated overpayment.

**Random Number Generator**

The selection of the random sample begins with the production of a random number between 0 and 1. This interval is inclusive of 0, but not inclusive of 1. It is commonly denoted $[0, 1)$. The random number generator used by OPI employs the same method used by the Office of the Inspector General, U.S. Department of Health and Human Services in its RAT-STATS software package and by Microsoft in Excel 2003.

The random number in $[0, 1)$ is generated by taking the fractional sum of three random numbers. The generation of the three initial random numbers begins with three seeds $x_0$, $y_0$, and $z_0$ that are between 1 and 30,000. The second set of seeds is based on the initial seeds as follows:

$$x_1 = \text{mod}(171x_0, 30269);$$

$$y_1 = \text{mod}(172y_0, 30307);$$

and

$$z_1 = \text{mod}(170z_0, 30323).$$

The function “mod” is the modulus of the two arguments that follow.

The first random number is the fractional part of $x_1/30269 + y_1/30307 + z_1/30323$. The next set of seeds, $x_2$, $y_2$, and $z_2$ are constructed from $x_1$, $y_1$, and $z_1$, respectively, in the same manner that $x_1$, $y_1$, and $z_1$ are constructed from $x_0$, $y_0$, and $z_0$. This process is continued iteratively.

The initial seeds are constructed from the provider ID (7 or 9 digit numeric field), the audit year (4 digits), and a counter (two digits). The counter will accommodate the possibility of taking more than one sample from a particular provider in an audit year. The counter for the first audit is “01”. If the provider ID is 7 digits in length, then it is concatenated with two zeroes at the beginning to form a 9 digit numeric value. These leading zeroes are not required when the provider ID’s length is 9 digits. The provider ID is then defined as $p_1p_2p_3p_4p_5p_6p_7p_8p_9$, the
audit year is defined as \( y_1y_2y_3y_4 \), and the counter is defined as \( c_1c_2 \). The initial seeds are then defined as the following integers of varying lengths depending on leading zeroes:

\[
\begin{align*}
x &= \text{value}(p_1p_2p_3p_4p_5p_6p_7p_8p_9y_1y_2y_3y_4c_1c_2); \\
y &= \text{value}(p_3p_4p_5p_6p_7p_8p_9y_1y_2y_3y_4c_1c_2p_1p_2); \text{ and} \\
z &= \text{value}(p_5p_6p_7p_8p_9y_1y_2y_3y_4c_1c_2p_1p_2p_3p_4).
\end{align*}
\]

From these, integers less than 30,269, 30,307, and 30,323 are created:

\[
\begin{align*}
x' &= \text{mod}(x, 30269); \\
y' &= \text{mod}(y, 30307); \text{ and} \\
z' &= \text{mod}(z, 30323).
\end{align*}
\]

If one of these is zero, then that variable is then set to the maximum of the other two.

Since the initial seeds should be less than or equal to 30,000, the seeds are adjusted as follows:

\[
\begin{align*}
x_0 &= \text{round}(x' \times 30000 / 30269); \\
y_0 &= \text{round}(y' \times 30000 / 30307); \text{ and} \\
z_0 &= \text{round}(z' \times 30000 / 30323).
\end{align*}
\]

**Sample Selection**

Prior to sample selection, the claim lines within each stratum are sorted by recipient ID (ascending order), TCN (ascending), and line number (ascending). The first random number in \([0, 1)\) generated is defined as \( r_1 \) and the size of stratum 1 is defined as \( N_1 \). The claim lines are then numbered 1 to \( N_1 \). The first claim line selected for sampling is \( 1 + \text{floor}(r_1 \times N_1) \). The function “floor” removes the fractional part of the number.

Once the first claim line is selected, it is removed from the list and the process is repeated in stratum 1 until the desired number of claim lines has been sampled. The process is repeated for
each of the 20 strata. It should be noted that the seeds for the random number generator are not reinitialized for each stratum.

**Difference of Means Test**

A two-tailed difference of means test based on the Normal distribution is performed at the 95% confidence level between the sample reported values and the overall population of reported values. The purpose of this test is to identify samples that may not be representative of the population from which they were sampled. If the sample fails the difference of means test, then the sampling procedure is repeated until the sample no longer fails the test. The statistic for this test is the “z-statistic of the sample mean”. That is, re-sampling should occur if the test statistic is less than -1.96 or greater than 1.96.

The sample mean is calculated using the formula: $\bar{x} = \frac{\sum_{i=1}^{20} N_i \bar{x}_i}{N}$. The variable $N_i$ is defined as the total number of claims in stratum $i$. The variable $\bar{x}_i$ is defined as the average claim amount for the sampled claims in stratum $i$ and $N$ is defined as the total number of claims in all twenty strata.

Note that $\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$ where $x_{ij}$ is sampled item $j$ from stratum $i$ and $n_i$ is the sample size for stratum $i$. The formula used for calculating the z-statistic is given by $z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n - 1}}$. The variable $\mu$ is the population mean, $\sigma$ is the population standard deviation, and $n$ is the overall sample size for all twenty strata.

**L&E Assessment of Sample Method**

**Stratification versus Non-stratification**

The choice to use a stratified sampling technique appears statistically reasonable and statistically acceptable since the reported amount of a claim is highly correlated with the variable that is being investigated (i.e., the correct claim amount). Stratification typically produces a more
precise point estimate of the correct total claim amount than simple random sampling. That is, the stratification process reduces variation versus simple random sampling.

**Number of Strata**

Cochran examined several populations and concluded that the benefit of stratifying the sample eventually hit a point of diminishing returns. Cochran estimated that diminishing returns often occur around 6 strata. However, the work of Neter and Loebbecke\(^3\) indicates that sampling precision generally increases as the number of strata increases. Their work also stated that there are cases where 20 strata produce results more precise than 15 strata.

It should be noted that a point of diminishing returns does not imply that the estimates become worse when a larger number of strata are used. Rather, it implies that the gains from increasing the number of strata eventually become negligible.

Even though the 20 strata used by OPI may be higher than necessary, there is no evidence that using 20 strata produces worse results than if a smaller number of strata were used. Therefore, OPI’s choice to use 20 strata for extrapolation appears statistically reasonable and statistically acceptable.

**Stratification by Dollar Amounts**

When there is a variable \(x\) that is highly correlated with the variable being investigated, then the variable \(x\) is typically considered a good candidate for defining strata. Since the reported claim amount is highly correlated with the correct claim amount, stratifying by book amount is common practice in financial auditing.

However, creating the strata by another correlated variable, such as procedure code, could be considered. Since OPI audits a myriad of provider types, a separate analysis of procedure codes by provider type would be required. Since placing claims with the same procedure codes in the

same strata will typically place claims with similar reported values in the same strata, there is likely no benefit to the significant expense in stratifying based on procedure codes.

OPI’s choice to stratify by dollar amount appears statistically reasonable and statistically acceptable.

**Strata Boundaries**

The method employed by OPI to define strata boundaries follows a method presented by Cochran (pp. 127-131) and also employed by Neter and Loebbecke (pp. 78-79) and Beck\(^4\) (p. 20) in their research. OPI employs an additional stage to the strata construction process that is similar to one alluded to by Cochran. This additional stage takes place when the trial strata are further subdivided into detail strata so that no detail strata contain more than 3% of the claims. When the subdivision of a trial stratum into detail stratum is triggered (because the trial stratum contains more than 3% of the claims), the method employed by OPI attempts to place an equal number of claims in each detail stratum.

**Sample Size**

The research cited to support the combined stratified regression interval estimator indicates that a sample size of at least 200 will produce statistically valid results. In the simulations performed by Schwarz\(^5\), 192 claims were sampled each time. The results referenced in Beck were based on samples of 200. Therefore, when utilizing this estimator, OPI should use a sample size of at least 200.

There are several methods that may be employed to estimate the sample size needed to achieve the desired level of precision. One possible approach is to use a formula to estimate the required sample size. Such a formula would depend on an estimate of the variance of the audit amount of each claim. In theory, the book amount of each claim could be used as a proxy in this


calculation. However, in practice, the variance of book amounts has typically underestimated the variance of correct or audit amounts. The apparent reason for this is that the audit amounts tend to be either identical to the book amounts or zero (that is, denied).

Another approach is to take a test or probe sample. This is commonly done when the variance of the population being sampled is unknown. However, due to the large number of audits conducted by OPI over the years, experience may be a stronger indicator of the required sample size than even a probe sample. Using experience to indicate an appropriate sample size is an acceptable approach.

To test the reasonableness of the OPI’s minimum sample size of 300, L&E utilized a set of statistical formulas that assumes a Normal distribution for large populations:\(^6\): 

\[ n_0 = \frac{Z^2 \times p \times (1-p)}{c^2} \quad \text{and} \quad n = \frac{n_0}{1 + \frac{(n_0-1)}{N}} \]

\(Z^2\) is defined as the abscissa of the normal curve that is based on a desired level of confidence. An assumed error rate of \(\alpha\) would produce a \(Z\) with the desired confidence level of 1- \(\alpha\) (e.g., an \(\alpha\) of 0.05 would produce an estimate with a confidence level of 95%). The variable \(c\) is the desired level of precision. The variable \(p\) is the estimated proportion of an attribute that is present in the population. \(N\) is the size of the population.

Because a proportion of 0.5 indicates the maximum variability in a population, it was used in evaluating the appropriateness of the method used by OPI. This approach is considered conservative. That is, the calculated sample size may be larger than when the true variability of the population is used.

When considering the application of the minimum sample size, the minimum sample size of 300 overrides the calculated sample size. That is, the precision achieved employing the OPI methodology is much better than 5% at a 95% confidence level.

Chapter 3
Extrapolation of Audit Results

INTRODUCTION

Once the sample size for each stratum is determined, the claims to be sampled for the audit are chosen from each stratum. The results of the claim audits are then extrapolated to the entire population of claims for each provider to produce a “fair” estimate of the total correct amount that should have been paid to the provider.

There are multiple ways to statistically extrapolate audit results by determining the “best” estimate based on a simple random sample.

Estimators

The extrapolation of audit results may or may not reference the original reported amount. A method of estimation that does not reference the original reported claim amount is called the mean-per-unit method. For example, suppose that the average audited claim amount from the sample was $100. Further, suppose that the total number of claims was 1,000. Then the best estimate for the total correct aggregate claim amount using the mean-per-unit method would be $1,000 \times $100 = $100,000.

Even though the mean-per-unit method generates an unbiased estimate of the total correct claim amount, methods that reference the original reported amount are more precise.

Estimators that take advantage of any correlation between the amount that is being estimated and any additional piece of information are called auxiliary information estimators, or, simply, auxiliary estimators. OPI uses the original reported amounts as auxiliary estimators.
There are three primary auxiliary estimator methods used in analyzing the relationship between reported and audited costs:

1. **Difference Estimator** - Examines the difference between the reported and audited cost.
   - That is, the dollar amount disallowed is analyzed.

2. **Ratio Estimator** - Examines the percentage difference between the reported and audited cost.
   - That is, the amount disallowed divided by the reported cost is analyzed.

3. **Regression Estimator** – Examines the linear relationship that may exist between the reported and audited costs.
   - That is, the line that minimizes the variation in reported and audited claim amounts is analyzed.

Method 1 is typically used if there appears to be a common disallowance figure. That is, the audited cost is some dollar amount (e.g., $20,000) less than the reported cost regardless of the magnitude of the reported cost.

Method 2 is typically used if there appears to be a consistent percentage reduction, such as the audited cost is consistently 97.5% of the reported cost.

Method 3 is typically used when the auxiliary variable (e.g., reported cost) is linearly related to the estimated variable (e.g., audited cost) but does not pass through the origin. That is, when the costs are linearly related but not expected to be zero.

It should be noted that if the average reported value of the sample is the same as the average reported value of the entire population of claims, then the auxiliary estimators would extrapolate to the same result as the mean-per-unit estimator. That is, the auxiliary estimators can be thought of as including a correction factor when there is a difference between the average sample reported value and the average population reported value.
**Fair Estimators**

The methods above describe ways to statistically arrive at the “best” estimate for the total correct claim amount. When sample results are extrapolated to an entire population, a “best” estimate is typically presented along with upper and lower limits associated with the estimate.

For example, a typical extrapolation statement might be: “The total correct amount that should be paid is $100,000 plus or minus $10,000, with a confidence level of 95%.” This statement means that with 95% confidence, the total correct claim amount is between $90,000 and $110,000. However, when corrected amounts are estimated to be used in the calculation of recoveries from a provider, a single point estimate must be utilized.

In the example above, the obvious choice for a point estimate would be $100,000 since it was calculated to be the “best” estimate of the total correct claim amount; however, rather than using the “best” estimate, it is common in audits to calculate an estimate that is considered “fair”. The “best” estimate would be increased or decreased such that it would be favorable to the provider being audited.

A common approach is to use a 95% confidence level to determine a “fair” estimate. That is, if there is a 95% confidence that the total correct claim amount is less than a certain value, then that value is the “fair” estimate.

In the above example, the “fair” estimate would be greater than the $100,000 best estimate. Since there is a 95% confidence that the total correct claim amount is between $90,000 and $110,000, there is a larger than 95% confidence that the total correct claim amount is below $110,000. Therefore, the “fair” estimate would be somewhere between $100,000 and $110,000 in order to reduce the risk that the provider is underpaid.

**OPI Extrapolation Method**

OPI employs a method that uses a combined stratified regression estimator and a Normal approximation with a 95% confidence limit. The approach creates a “fair” estimate in which the risk of underpayment is 5%.
The variable \( n_i \) is defined as the number of claims sampled from stratum \( i \). Variables \( x_{ij} \) and \( y_{ij} \) are the reported and correct (or audited) amounts, respectively, for the \( j^{th} \) claim from stratum \( i \).

The covariance between \( y \) and \( x \) in stratum \( i \) is given by \( S_{yx_i} = \frac{\sum(y_{ij} - \bar{y}_i)(x_{ij} - \bar{x}_i)}{N_i} \), where the summation is over all claims in stratum \( i \). Recall that \( \bar{x}_i \) and \( \bar{y}_i \) are the average reported and correct amounts, respectively, for all claims from stratum \( i \). Variables \( \bar{x}_i \) and \( \bar{y}_i \) are the average reported and correct amounts, respectively, for all claims sampled from stratum \( i \). The estimate of \( S_{yx_i} \) used in the calculations is \( s_{yx_i} = \frac{\sum(y_{ij} - \bar{y}_i)(x_{ij} - \bar{x}_i)}{n_{i-1}} \) where the summation is over all claims sampled from stratum \( i \).

Similarly, estimates for the variance of the correct and reported amounts are given by

\[
s_{y_i}^2 = \frac{\sum(y_{ij} - \bar{y}_i)^2}{n_{i-1}}, \text{ and } s_{x_i}^2 = \frac{\sum(x_{ij} - \bar{x}_i)^2}{n_{i-1}}. \text{ The summations are over all claims sampled from stratum } i.
\]

The combined linear regression estimate of the correct amount is given by \( \hat{y} = y_{st} + b(X - x_{st}) \), where \( y_{st} = \sum N_i \bar{y}_i \), \( x_{st} = \sum N_i \bar{x}_i \) and \( X \) is the (known) sum of all reported amounts for the population of claims. The value of \( b \) that minimizes the variance of the regression estimator is given by

\[
b = \frac{\sum(N_i(N_i - n_i) s_{yx_i})}{\sum(N_i(N_i - n_i) s_{x_i}^2)}
\]

where both summations are over all strata.

An estimate of the variance of the estimator \( \hat{y} \) is given by

\[
\nu(\hat{y}) = \sum N_i(N_i - n_i) \left( s_{y_i}^2 - 2bs_{yx_i} + b^2s_{x_i}^2 \right).
\]

The following formulas used to calculate the upper precision limit of the combined linear regression estimate are presented without justification. The upper precision limit of the correct value is defined as \( \hat{y} + 1.64485\sqrt{\nu(\hat{y})} \).
The total amount that the provider owes the state will also include:

1. The difference in the audited and reported total amounts for the 25 highest dollar claims; and
2. The difference between the amount reimbursed to the provider and the upper precision limit.

This amount is \( X - [\hat{y} + 1.64485\sqrt{v(\hat{y})}] \).

**L&E ASSESSMENT OF EXTRAPOLATION METHOD**

**Combined Regression Estimator**

Auxiliary information estimators (which include the regression estimator) are generally considered the most precise audit sampling estimates\(^7\).

Additionally, the regression estimator is potentially more precise than the other commonly used auxiliary information estimators (difference and ratio estimators). This occurs when the relationship between two variables is approximately linear, but it does not go through the origin. In these situations, an estimate based on the regression of one variable on the other is preferred, rather than an estimate based on the ratio of the two variables\(^8\).

Therefore, the OPI’s selection of a combined regression estimator appears statistically reasonable and statistically appropriate.

**Upper-Limit Estimate**

The distributions of errors in audit populations can be highly skewed. Due to this characteristic, these types of populations may require larger than typical sample sizes for Normality assumptions to hold. Additionally, audit populations can vary widely with respect to both the

---


error rate and the types of errors that could occur (e.g., overstatement errors, understatement errors, or a combination of the two).

Despite the potentially skewed nature of audit populations, research provides consistent support for the validity of OPI’s use of the upper limit of a one-sided 95% confidence interval based on the Normal distribution. Tamura\(^9\), Beck, Neter and Loebbecke, and Schwarz all indicate that the upper limit of a one-sided 95% confidence interval based on Normality will be overly conservative in favor of the medical provider.

Schwarz is of particular interest since his research directly related to OPI populations.

As a result, OPI’s selection of an upper-level estimate appears statistically reasonable and statistically appropriate.

---

Chapter 4

Conclusion

GENERAL ASSESSMENT

Lewis & Ellis, Inc. (L&E) was engaged by OPI to assess the appropriateness and validity of the “combined stratified regression interval estimator”, which is the primary statistical approach utilized by OPI. Pursuant to WAC 182-502A-0900, OPI may employ other statistical methods in certain situations. Those methods were not evaluated.

Cochran is a standard graduate level text on the subject of sampling techniques. The methods and results found in Cochran provide the theoretical foundation for the method employed by OPI.

The work of Schwarz examined 4,232 actual hospital claims audited by OPI. Schwarz contains empirical results involving audit populations and populations constructed to be representative of audit populations. Schwarz concluded that the one-sided interval estimator is conservative (relative to a normally distributed estimator) in favor of the provider. The results of Schwarz provide evidence of the efficacy of the method employed by OPI that is based on the statistical techniques of Cochran.

All of the research applicable to the process employed by OPI consistently supports the conclusion that the regression estimator and the normal approximation of the one-sided confidence limit are slightly conservative to the benefit of the provider.
Therefore, based on the analysis performed:

- The OPI approach to statistical auditing appears reasonable and appropriate;
- The OPI approach to statistical auditing should meet and exceed industry standards; and
- There is no research available to indicate that an alternative approach would consistently produce better results.

**CONSISTENCY WITH STATUTORY AUTHORITY**

The OPI is responsible for maintaining an ongoing program to audit providers participating in HCA programs. This responsibility and authority is governed by multiple Federal and State laws and regulations.

The following 6 sections specifically relate to audit sampling, extrapolation, and claim review.

**Section (1)**

The department's procedures for auditing providers may include, but are not limited to, the following:

(a) The use of random sampling and extrapolation; and/or

(b) A claim-by-claim based review.

*Based on discussions with OPI and the analysis performed, OPI appears to be following this section of WAC 182-502A-0900.*
**Section (2)**
The department's sample sizes are sufficient to ensure a minimum of ninety-five percent confidence level.

(a) When calculating the amount to be recovered, the department totals all overpayments and underpayments reflected in the sample and may extrapolate to the universe from which the sample was drawn.

(b) When the department uses the results of an audit sample to extrapolate the amount to be recovered, the provider may request a description of all of the following:
   (i) The universe from which the department drew the sample;
   (ii) The sample size and method that the department used to select the sample; and
   (iii) The formulas and calculation procedures the department used to determine the amount of the overpayment.

(c) If a provider rebills a claim(s) for an adjustment and that claim(s) is part of the audit universe, the department does not remove the original paid claim(s) amount from the audit universe

*Based on discussions with OPI and the analysis performed, OPI appears to be following this section of WAC 182-502A-0900.*

**Section (3)**
When a claim-by-claim audit is conducted, specific claims are selected from the universe and audit overpayments are not extrapolated.

*Based on discussions with OPI and the analysis performed, OPI appears to be following this section of WAC 182-502A-0900.*

**Section (4)**
The department recovers overpayments identified in the final audit report.

*This section was not applicable to the review.*
Section (5)
The department does not consider non-billed or zero paid services or supplies when calculating underpayments or overpayments.

*Based on discussions with OPI and the analysis performed, OPI appears to be following this section of WAC 182-502A-0900.*

Section (6)
The department considers undocumented services to be program overpayments.

*This section was not applicable to the review.*
Appendix A

ASOP 41 Disclosures

The Actuarial Standards Board (ASB), vested by the U.S.-based actuarial organizations\(^{10}\), promulgates actuarial standards of practice (ASOPs) for use by actuaries when providing professional services in the United States. Each of these organizations requires its members, through its Code of Professional Conduct\(^{11}\), to observe the ASOPs of the ASB when practicing in the United States. ASOP 41 provides guidance to actuaries with respect to actuarial communications and requires certain disclosures which are contained in the following.

**Identification of the Responsible Actuary**

The responsible actuary is David M. Dillon, FSA, MAAA, MS, Vice President & Principal at Lewis & Ellis, Inc. (L&E). This actuary is available to provide supplementary information and explanation. The actuary also acknowledges that he may be acting as an advocate.

**Identification of Actuarial Documents**

The date of this document is January 24, 2014. The date (a.k.a. “latest information date”) through which data or other information has been considered in performing this analysis is January 6, 2014.

**Disclosures in Actuarial Reports**

- The contents of this report are intended for the use of the Office of Program Integrity (OPI), within the Washington Health Care Authority (HCA). The limitations on the use

\(^{10}\) The American Academy of Actuaries (Academy), the American Society of Pension Professionals and Actuaries, the Casualty Actuarial Society, the Conference of Consulting Actuaries, and the Society of Actuaries.

\(^{11}\) These organizations adopted identical *Codes of Professional Conduct* effective January 1, 2001.
or applicability of the actuarial findings are that it is intended for internal documentation for OPI and these communications should not be relied upon for any other purpose.

- Lewis & Ellis Inc. and Healthcare Actuaries are financially and organizationally independent from the providers whose claim audits were analyzed. There is nothing that would impair or seem to impair the objectivity of the work.

- The purpose of this report was to provide the OPI an assessment of the statistical audit procedures under the Department’s regulatory purview.

- The responsible actuary identified above is qualified as specified in the *Qualification Standards* of the American Academy of Actuaries.

- Lewis & Ellis reviewed the statistical audit procedures based on the data, files, communications, and documents provided by OPI. L&E nor the responsible actuary assumes responsibility for these items that may have a material impact on the analysis. We have reviewed the data for reasonableness, but we have not audited it. To the extent that there are material inaccuracies in, misrepresentations in, or lack of adequate disclosure by the data, the results may be accordingly affected.

- We are not aware of any subsequent events that may have a material effect on the findings.

- There are no other documents or files that accompany this report.

- The findings of this report are enclosed herein.

**Actuarial Findings**

The actuarial findings of the report can be found in the body of this report, as well as the attached exhibits.

**Methods, Procedures, Assumptions, and Data**

The methods, procedures, assumptions and data used by the actuary can be found in body of this report.
Assumptions or Methods Prescribed by Law

This report was prepared as prescribed by applicable law, statutes, regulations and other legally binding authority.

Responsibility for Assumptions and Methods

The actuary does not disclaim responsibility for material assumptions or methods.

Deviation from the Guidance of an ASOP

The actuary has not deviated materially from the guidance set forth in an applicable ASOP.