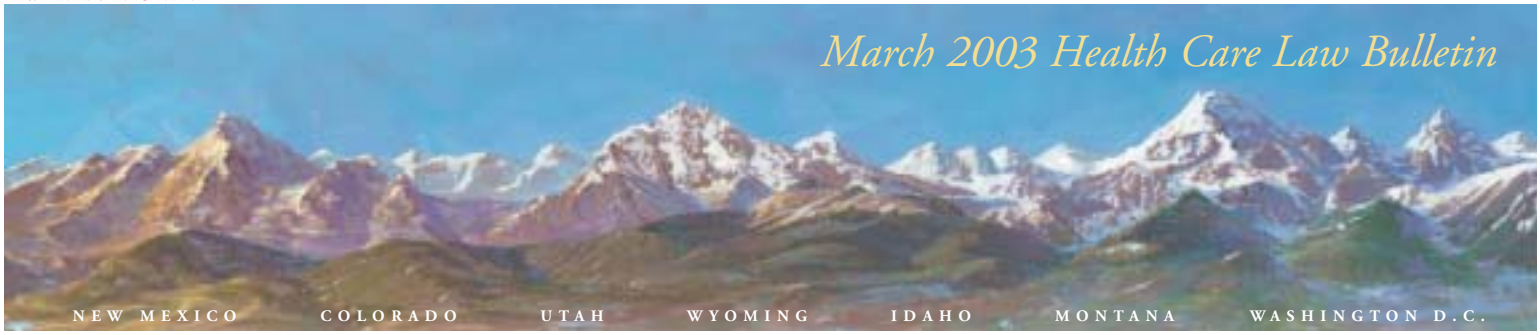

Rocky Mountain Health Care

HOLLAND & HART LLP
The Law Out West℠
www.hollandhart.com

Roaming the Random Range

Defending against governmental extrapolation of high dollar recoupment claims through statistical sampling in Medicare and Medicaid overpayment and fraud actions.

Executive Summary

When federal or state Medicare or Medicaid oversight agencies use "statistical sampling" techniques to extrapolate overpayment amounts from a universe of claims based on a limited sample—human process errors can provide the foundation for an effective defense.

Statistical Sampling

It can be a shock to open a government overpayment determination notice and view the announcement of a recoupment action against your company to reclaim say $3,000,000.00 in alleged overpayments based on an agency's "statistical sampling" methodology. Welcome to the bewildering world of statistical sampling and extrapolation of data in the public funding of health care services. The government has taken a "footprint" review of a few of your files, deemed your documentation wanting and has extrapolated a "dinosaur" determination that the universe of your submitted claims is similarly wanting and your company owes big bucks and maybe even the "farm" to the government.

The agency formerly known as the Health Care Finance Administration[1] (HCFA) adopted a rule permitting the use of statistical sampling techniques as part of the armamentarium of the government in the exercise of its program review and integrity function. The Social Security Act requires the government to review, identify and/or deny inappropriate, medically unnecessary, excessive, or routine services.[2] Sampling may be used where claim volume of a provider under review is "voluminous," the claims reflect a "pattern of overbilling," and a case by case review is "not administratively feasible."[3]

Given the large number of claims processed by home health agencies and other health care providers relying on some level of federal reimbursement, the use of statistical sampling in lieu of a review of all of a provider's claim files is a cost saving boon to federal and state governments administering Medicare and Medicaid programs. It also has an extraordinary "in terrorem" effect on providers because of the process of extrapolation of small review samples into huge financial obligations cutting across all claims submitted during the audit period.

Home health agencies have been particularly vulnerable to statistical sampling problems because of the complexity and difficulty of maintaining adequate file documentation in a labor intensive

CONTINUED ON PAGE 2

Defending Against High Dollar Recoupment Claims through Statistical Sampling

CONTINUED FROM PAGE 1

enterprise performed in patient homes. The government treats all services not adequately documented as not having been provided, and therefore the basis of an overpayment claim and recoupment action.

A typical audit usually starts with a "random" selection and review of a small number of files to determine the adequacy of the documentation for claims previously presented for payment. The files, once identified by a computer "randomizing" program, are reviewed by nurses or other trained personnel representing the government to see if there exists a "pattern" of overbilling. If the government perceives a pattern, it will select and review a larger randomized sample of claims (*i.e.* 100). The documentation error rate in the larger sample is then determined. If the hapless provider has another 10,000 claims during the audit period, a computer program is used to extrapolate the error rate in the 100 files over the entire 10,115 files and pretty soon we are talking about big money. The computer program, at the end of the process, usually spits out a high and low range of probable overpayment and the government usually selects the lower number (call it $2,900,000 instead of $3,000,000) just to show how conservative and careful the government is being. Your company has just been mortally wounded if not actually killed by a computer in the hands of government statisticians.

Ironically, the very agency cost concerns that led to the adoption of statistical sampling techniques also provide the seeds of defenses against them.

The seeming irrefutability of computer-generated numbers can be at least as terrifying as the size of the numbers generated through sampling techniques. To the uninitiated (which includes most of us) the nature and practice of statistical analysis is arcane and impenetrable. How does one defend against the clinical determinism of computer driven mathematics?

The answer is—human process flaws. Ironically, the very agency cost concerns that led to the adoption of statistical sampling techniques also provide the seeds of defenses against them. The reality is that there are significant costs in utilizing statistical sampling correctly and that review agencies, either through lack of funding or ignorance with the technical requirements of the sampling process, rarely get it right—leaving room for significant challenges to the validity and accuracy of the final numbers.

Lines of Defense

The first line of defense is in any inaccuracy of the initial review. If the claims reviewers made subjective judgments about the adequacy of the documentation in the claims files or if the provider is able to supplement missing data to the manually reviewed documentation, the validity of the deficient samples can be severely compromised under GIGO, the axiom of "garbage in-garbage out."

The second line of defense relates to "due process" or procedural fairness requirements for the validity and accuracy of the methodology used by the agency in determining the sample to be extrapolated. This is not the same as being "processed duly." The agency utilizing statistical sampling techniques has the burden to establish that the sample developed is in fact random and statistically valid. In the seminal case of Chaves County Home Health Services v. Sullivan, 931 F.2d 914 (D.C. Cir. 1991), the District of Columbia Circuit Court of Appeals held that the use of statistical sampling techniques was not in and of itself a violation of due process of law "in light of fairly low risk of error so long as the extrapolation is made from a representative sample and is statistically significant."[4]

Sample Size and Reproducibility

Is the sample truly representative and is the extrapolation statistically significant? There are two general sources of guidance as to the assurance of the representative accuracy of the sample and the statistical significance of the extrapolation, and both really relate to the precision in the tolerances of the sample.

The first source emanates from the rules and procedures adopted by the government. The second derives from generally accepted standards within those disciplines regularly engaged in the practice of statistical analysis. HCFA originally published its own Sampling Guidelines Appendix (SGA) in the Medicare Carrier's Manual setting out minimum standards to assure the integrity of the sample. The SGA identified the basic sampling unit, "a service, a bill, or a beneficiary for a particular period of time."[5] The SGA identified a number of factors affecting the accuracy of the sample—the time frame of the sample, the size of the sample, the size of the claim amount sought, the stratification of the sample universe, the randomness of the selection, and the complete documentation of the process so as to enable others to reproduce the results. The SGA explicitly recognized that the second source,—"persons with competence in statistical sampling can provide effective guidance in using more sophisticated techniques which might ensure a better result for the same degree of effort."[6]

SGA's Useful Sampling References

- W. G. Cochran, *Sampling Techniques*, 2nd Edition, New York, John Wiley and Sons, 1963
- Morris H. Hansen, William W. Hurwits, and William G. Madow, *Sample Methods and Theory*, New York, John Wiley and Sons, 1953
- Leslie Kesh, *Survey Sampling*, New York, John Wiley and Sons, 1965

The effect of compliance with the SGA "minimum standards" was to assure that certain precision standards in the results were achieved. Unfortunately, in practice, oversight agencies tended to ignore the standards or to farm out the sampling process to subcontractors who were unfamiliar with them. This was particularly true with respect to the selection of sample size to be used. Instead of using statistically significant sample sizes to achieve acceptable accuracy tolerances in the result, there was a tendency to select an arbitrary number of say 100 or 200 when the SGA and other authorities might require a minimum of 400. (The actual minimum number required can be determined mathematically and almost

always ends up being an odd number like 353 rather than a round number like 200, which is an indication of too low an arbitrary number used because of agency resource limitations. See *In the Case of American Health Care Services,* HICN 103-01-0077A (2000) before the Social Security Administration, Office of Hearings and Appeals overturning a $1,248,747 overpayment determination.) There was a representative of the Office of Inspector General who testified that the OIG could not have looked at the required minimum of 400-sample size due to "lack of audit resources and availability of staff" and the OIG failed to preserve the sample "frames" and other data to permit a replication of the sampling.

Some states have adopted their own standards for statistical sampling, while many have not. The applicability of the federal Medicare standards to state Medicaid recovery actions is not anywhere made explicit, but the logic of their use in state actions flows from the fact that both programs involve the recovery of federal funds.

In 2001, HCFA replaced the SGA with PMB-01-01[7] that eliminated the minimum sample size and sampling detail requirements contained in the SGA. It also suggests that the probability sample and statement results are "always" valid which is unsupported in any professional literature on the subject. This is a prime example of the principle "if you can't win by playing by the rules (or even not by the rules) just get rid of the rules." The fundamental problem for the government is that the procedure used by the agency must still stand up to due process requirements in order to be upheld, and the effect of this dilution of precision in the sampling requirements has yet to be determined in the courts.

Removing Bias from Samples

Sample size and reproducibility are two factors affecting the accuracy of a representative sample. A representative sample is one from which all bias has been removed. A basic sample is random if every name or element in the whole group has a mathematically equal chance to be in the sample. The question is how accurate a sample can be taken to represent the whole universe measured in figures (i.e. "probable" error and "standard error").

The degree to which the sample universe is homogenous is an important accuracy factor. The greater the degree of heterogeneity the more difficult and complicated the process. In a recent audit in Colorado, the state Department of Health Care Policy and Finance utilized sample units described as TCNs (Transaction Control Numbers). These were individual billings for one patient for varying periods, encompassing different units of service. No effort was made to use more homogenous units or to "stratify" the disparate elements of the units into discrete categories so as to establish greater reliability in the sample accuracy. Further, the states predicated its analysis on "rows of TCNs." It discarded entire rows of units of service when there was any defect in the documentation of any individual TCN in the row. There were also two huge "outliers" in the selected samples which alone accounted for twenty percent of the sample claims. The final result was an asymmetrical distribution of the sample. (A normal distribution looks like a bell curve with the mean being the same value as the median.) The states mean of $100.69 was asymmetrical from the median of $33.72, reflecting a demonstrable lack of precision in the integrity of sample as fairly representing the universe of TCNs.

The coefficient of variation (COV) is the best overall measure of the validity of the sample.

Government agencies will sometimes rely on concepts like the "central limit theorem" to compensate for the lack of stratification or homogeneity in the sample. The theorem provides that when averaging over an increasing number of different elements with varied distribution, averages of those elements become increasingly closer to normal or "Gausian" distribution—a standard distribution. The problem is that it takes a very large sample to reach the standard distribution and the argument is therefore circular.

It is amazing how infrequently agencies calculate the coefficient of variation (COV) of the sample which is a mathematical measure of the imprecision of the sample. The higher the value, the more imprecise the sample. The COV is the best overall measure of the validity of the sample. In order to achieve improvement in the tolerances as measured by the COV, the agency must adjust for outliers, stratify sample categories and/or increase the sample size.

Once statistically acceptable precision in the sample is determined, there are a number of methods of extrapolation that can be applied to reach a representative amount. The method selection is generally not statistically significant in recovery actions unless clerical errors exist.

Despite the perception of mathematical unassailability, claims of overpayment and fraud that are developed through statistical sampling are rarely done so with sufficient care and precision to overcome the basic constraints of due process of law and fundamental fairness. There is almost always room to develop a formidable defense in statistical sampling recoupment actions.

ENDNOTES

[1] Now the 'friendlier' sounding "Centers for Medicare Services" ("CMS").
[2] Section 1842 (a)(2)(6) of the Social Security Act. See also 42 C.F.R. §421.200.
[3] HCFA Rule 86-1.
[4] Chaves at 922.
[5] MCM, SGA §2.5.
[6] MCM, SGA §1.1
[7] See PIM, Exhibits 7-7.7.

HOLLAND & HART LLP
The Law Out West℠
www.hollandhart.com

Attorneys at Law
Health Care Practice Group
Suite 3200
555 Seventeenth Street
Denver, Colorado 80202-3979

How to Defend Against Government High Dollar Recoupment Claims

March 2003