BEFORE THE ALASKA OFFICE OF ADMINISTRATIVE HEARINGS ON REFERRAL BY THE COMMISSIONER OF HEALTH AND SOCIAL SERVICES

In the Matter of:	)	
	)	
C CARE SERVICES LLC.	)	OAH No. 11-0015-DHS
	)	

DECISION

I.   Introduction

A.   Summary

This is a recoupment case for alleged overcharges to the Medicaid program. C Care Services LLC offers personal care assistance and home care services to elderly and disabled Alaskans who are on Medicaid. The Department of Health and Social Services, Division of Health Care Services (Division) audited C Care's billings to the program for the calendar year 2007, during which C Care submitted just over 14,000 claims and received approximately $3.97 million from the program. Based on this audit, the Division presently contends that it was overcharged by just under $50,000, of which about $40,000 has not yet been reimbursed.

The sole matter still in dispute is the validity of the statistical sampling and extrapolation supporting the main component of this overcharge claim. Through a combination of exchanges of written testimony and live cross-examination, the parties tried that issue before the undersigned between July and November, 2011. The primary witnesses were Ted Eschenbach, PhD, an emeritus professor at the University of Alaska-Anchorage, and Dale Lehman, PhD, currently a professor at Alaska Pacific University. Both have extensive training in statistical theory and have taught allied subjects.

The evidence showed that the statistical methodology used by the auditor was imperfect. Nonetheless, using the sample collected, the Division was able to demonstrate to a reasonable certainty, by means of a statistically valid methodology, that an overpayment occurred in an amount equal to or greater than the Division's claim.

B.   Procedural History and Amounts at Issue

This case grows out of an audit completed on October 12, 2010 by Myers and Stauffer LC, under contract to the Division. Largely by means of sampling and extrapolation, the audit generated an overpayment amount of $495,486.22.[1] After adjusting for prior repayments by

---

[1]   Agency Record (A.R.) 13, 50-51. The sampling and extrapolation process accounted for $492,556.72 of this amount. About $3,000 in overpayments were identified by other means, but none of these other findings or the methods behind them will be at issue in this decision.

C Care, the Division made a written demand for repayment of $486,835.11 on October 19, 2010.[2] It offered C Care a full evidentiary hearing to contest the demand, and stipulated that should a hearing be requested the Division would have the burden of proving the amount owed.[3] C Care requested the hearing on November 18, 2010. In January of 2011 William Hogan, Commissioner of Health and Social Services, referred the matter to the Office of Administrative Hearings for a recommended decision.

After referral, the case was placed on a slow track while the parties exchanged additional documents and explored settlement. These efforts were largely successful, resulting in a revised audit report dated April 27, 2011.[4] The findings identified in the revised report, and their status in the case as ultimately presented for hearing, are summarized below:

	Finding	Alleged Overpayment	Status
A	Duplication of claims	$598.50	Not contested.
B	Conflicts with inpatient stays	$1,606.50	Not contested.
C	Conflicts with other pers. care agencies	$724.50	Not contested.
D	Errors in non-duplicate, non-conflict claims	$45,683.20	$897.75 not contested. Remainder contested.

The total of A through D, representing the Division's revised overpayment claim, is $48,612.70. Of this, $3827.25 is not contested, as reflected in the table, and that amount has already been reimbursed. Moreover, the parties agree that credit must be given for an additional $4,823.86 that C Care has already paid. In sum, the case to be decided is limited to Finding D from the table above, and the potential amounts owing are as follows:

■   If Finding D is correct, C Care owes $39,961.59 ($48,612.70 minus $3827.25 minus $4,823.86);

■   If Finding D is incorrect except as to the conceded $897.75, C Care has over-reimbursed $4,823.86.

---

[2]   A.R. 3.

[3]   *Id.* The hearing procedure offered to C Care was more protective of C Care's rights than the paper "appeal" provided for in former 7 AAC 43.1490 (in effect when the charges were incurred and when the audit was initiated) and in the substantially identical 7 AAC 160.130 (in effect when the audit was completed). Because it was offered, the more protective procedure needed to be followed. *See Hidden Heights Assisted Living, Inc. v. State, Dep't of Health & Soc. Serv.*, 222 P.3d 258, 269 (Alaska 2009).

[4]   The revised report is not part of the numbered agency record, but became part of the case record as Exhibit 6 to C Care's Opening Brief.

C. Evidence Received

The record for decision consists of the following items:

- Agency Record stamped 1-1493;

- Written testimony of Dale Lehman and C Care Exhibits 1-6 attached to C Care's "Opening Brief" of July 11, 2011;

- Written testimony of Allan Hansen attached to the Division's "Pre-Hearing Brief" of July 11, 2011.

- Written testimony of Ted Eschenbach filed by the Division on August 25, 2011;

- Oral testimony received on September 12, 2011;

- C Care Exhibits 7, 8, and 9 admitted on September 12, 2011;

- Additional written testimony of Ted Eschenbach admitted on September 12, 2011 (styled as a memorandum from Eschenbach to Friend dated September 11, 2011);

- Additional written testimony of Ted Eschenbach filed on September 26, 2011;

- Additional written testimony of Dale Lehman filed on September 26, 2011.

II. Regulatory Framework

Regulation 7 AAC 160.110[5] authorizes the Department of Health and Social Services to audit Medicaid providers. The regulation authorizes the department to gather information "sufficient to support a reasonable basis for determining the provider's compliance with the legal requirements of the Medicaid program."[6] A second regulation, 7 AAC 160.120,[7] provides that the "department or its designee may use *statistically valid sampling methodologies* to . . . calculate overpayment amounts" (italics added).

III. Facts

A. Myers and Stauffer Sample

The portion of the audit at issue in this case addressed the accuracy of 14,031 claims submitted by C Care, totaling just under $4 million. The average size of the claims in the audit pool was about $282.[8]

---

[5] A substantially identical regulation, former 7 AAC 43.1440, was in effect when the charges at issue were billed and when the audit was initiated.

[6] 7 AAC 160.110(i)(1) [former 7 AAC 43.1440(i)(1)].

[7] A substantially identical regulation, former 7 AAC 43.1470, was in effect when the charges at issue were billed and when the audit was initiated.

[8] Figures in paragraph drawn from or calculated from C Care Ex. 6. Note that the original universe was 14,054 claims, but 23 were removed (and disallowed) prior to sampling.

To evaluate this pool, Myers and Stauffer selected a random sample of 96 claims, representing $24,543.75 in payments. The average size of the claims in the sample was about $256.[9] The randomness and representativeness of the sample is not contested.

Among the 96 sampled claims, Myers and Stauffer identified 11 overpayments. The overpayments totaled $897.75. Averaged across the 96 claims in the sample, this represents an average overpayment per claim of **$9.35**, or 3.65%. The accuracy of the overpayment findings is not contested.[10]

B.   *Extrapolation*

Myers and Stauffer used this data to extrapolate a total overpayment across the universe of 14,031 claims. It calculated a one-sided 90% confidence interval for the average overpayment. A 90% confidence interval means that, if the full universe of claims were audited with perfect accuracy, there is a 90% chance that the true average overpayment would be greater than or equal to the lower bound of the interval.[11] The lower bound calculated by Meyers and Stauffer was an average overpayment per claim of **$3.26**,[12] or 1.27%.

Myers and Stauffer used the lower bound of the one-sided 90% confidence interval as the basis for its overpayment finding; that is, the overpayment finding represents $3.26 (not $9.35) multiplied across the universe of 14,031 claims. In seeking to develop a one-sided 90% confidence interval, Myers and Stauffer was applying a department audit policy. Insofar as it is in writing, the policy is not part of the record in this case.

It is customary in audit tests to use confidence intervals of 90% or higher,[13] but this is not universal. For medical provider audits in some jurisdictions, the midpoint of the sample is used, instead of the lower bound of a confidence interval.[14]

Using a confidence interval is a conservative approach, reducing the likelihood of an erroneously high overpayment finding.[15] Had the midpoint of the Myers and Stauffer sample been used, the calculated overpayment would have been approximately $131,000.[16]

---

[9]   *Id.*

[10]   *Id.*

[11]   Direct written testimony of Lehman.

[12]   Direct written testimony of Eschenbach; calculation from C Care Ex. 6. The exact figure seems to be 3.255876.

[13]   Sept. 12, 2011 written testimony of Eschenbach.

[14]   Oral testimony of Eschenbach. *See also, e.g., Michigan Dep't of Educ. v. U.S. Dep't of Educ.,* 875 F.2d 1196, 1199 (6th Cir. 1989) (federal audit; no confidence interval used despite very low sample sizes within strata).

[15]   *E.g.,* written testimony of Hansen, question 7.

[16] $9.35 x 14,031.

Myers and Stauffer calculated the confidence interval by a method that assumes that the real occurrence of errors in the universe of claims follows a "normal distribution."[17] In a normal distribution, the errors in the full universe of claims would graph into a bell curve centered on the mean.

C.   Problems with Methodology

The data from the 96-claim sample does not suggest a normal distribution at all. Instead, it suggests an error distribution in which the most common value, by far, is zero, with the frequency of errors then decreasing as the size of the error increases. The sample graphs as follows:
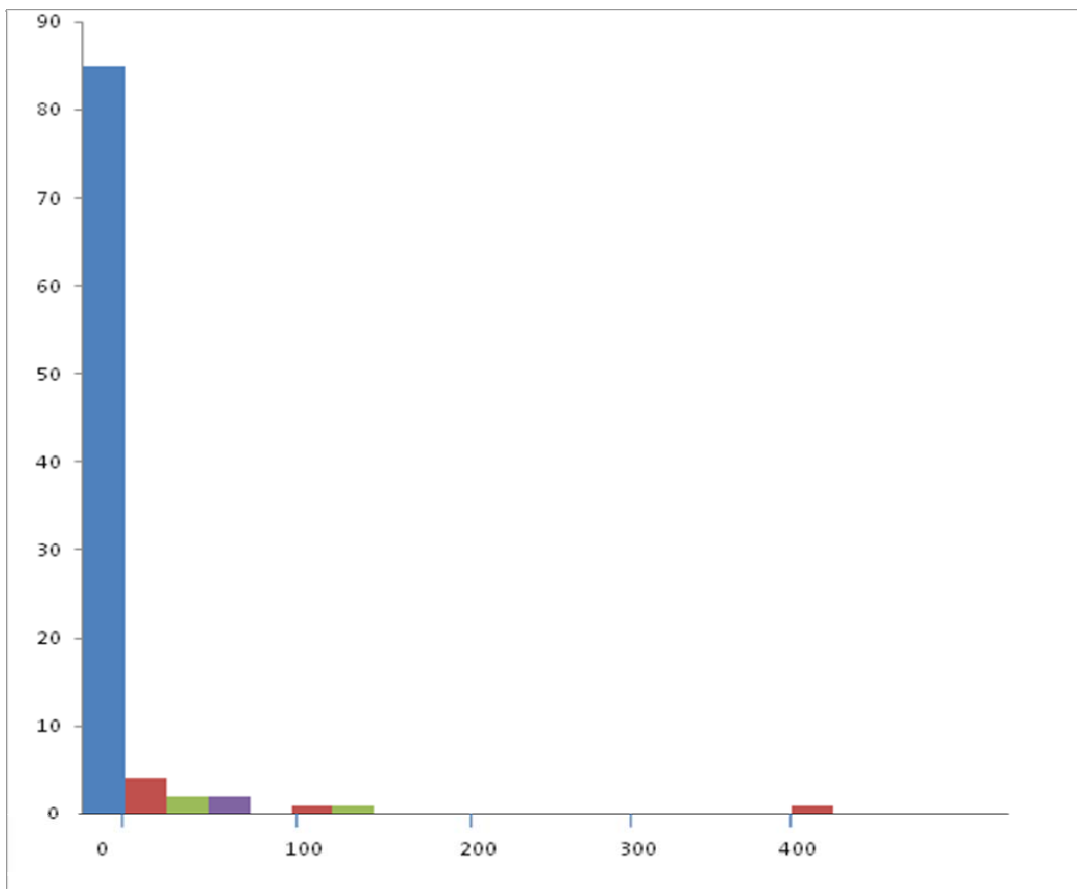
It will be helpful to make some intuitive observations about this data set before turning to statistical theory. First, an important feature of the sample is the single outlier on the right-hand margin. This corresponds to one overpayment of $409.50, drawn from a relatively large billing of $677.25 on claim D908709.[18] This one overbilling accounts for nearly half of the $897.75

---

[17]   Direct written testimony of Lehman; oral testimony of Eschenbach (assumption of normality underlies auditor's assumption that the T distribution can be used).
[18]   The total billing is found at A.R. 588. The audit finding is at C Care Ex. 6, Att. 2, p. 18.

total error that Myers and Stauffer identified in the sample set of 96 claims. As a matter of intuition, an extrapolation relying very heavily on one data point would need to be approached with some caution.

On the other hand, one must note that when it took a random slice of 96 claims out of the 14,031, Myers and Stauffer found 11 errors in that small slice, including three of more than $100. As a matter of common sense, we can be abidingly certain that there are other errors—indeed, many other errors—in the overall pool of 14,031. By way of analogy, imagine a thoroughly mixed drum known to contain 14,000 red and black marbles, but with an unknown percentage of each kind. If I reach blindly into the barrel 96 times and I pull out a red marble 11 of those times, I know, to a virtual certainty, that there are going to be more red marbles in that barrel. In other words, the likelihood that there were only 11 red marbles to begin with—scattered among 14,000—and I just happened to find them all the first 96 times I reached into the barrel, is utterly remote.

These two observations shape the more theoretical findings later in this decision. On the one hand, the intuitive risk of relying too much on a single outlier gives credence to the concerns of both parties' experts, both of whom were troubled by Myers and Stauffer's methodology and both of whom sought better ways to model the data. On the other hand, the second observation completely undermines the central opinion offered by C Care's expert. C Care's expert opined that the "only" extrapolation one can make from the data set *with 90% certainty* is that there were at least $897.75 in overpayments among the 14,031 claims.[19] This is equivalent to saying that, notwithstanding the significant number of errors we have encountered in the first 96 randomly-selected claims, there is a *greater than 10% chance* that the remaining 13,935 claims could be audited one by one without finding a single additional error. It is an untenable, almost preposterous position, and it discredits his analysis.[20]

D. Better Methodologies

The data points collected by Myers and Stauffer are strongly skewed to the right, that is, they graph with a long right-hand tail. This is not surprising in the context of an audit aimed solely at identifying whether payments made to the provider are supported by the service records

---

[19] Direct written testimony of Lehman at 9.
[20] The ALJ recognizes that Dr. Lehman shied away from making this assertion in so many words when the question was put to him directly during oral examination. Hearing audio file 3 at 25:30. However, he went on to say that he was opining on "what value am I 90% sure that the average overpayment exceeds this amount," or on the value for which "you can be 90% sure that this provider owes at least this amount." He stuck to his testimony that the correct answer to these questions is $897.75 in overpayments.

they rely on.[21] Values below zero would not ordinarily be possible (a service that was underbilled will be adequately supported and will simply show as a zero overpayment). One would expect a competent billing organization to have a large number of zero overbillings, and for the remainder of the data to cluster around small errors, with larger errors less and less common.

One way of approaching this data set, endorsed by both experts,[22] is to view it as two distributions rather than one. The first, with 96 data points, is concerned only with the frequency of overpayments (in this case, 11.5 % in the sample collected). The second, with 11 data points, is concerned with the size of overpayments when overpayments occur. This second sample still has a peak on the left with a long right-hand tail.

This kind of data is similar to what is seen with insurance claims or oil spill volumes, with a "long right-hand tail of relatively unlikely but extreme outcomes."[23] An approach that has been used over a long period to make sense of asymmetrical data of this kind is to assume, not a normal distribution, but instead a lognormal distribution.[24] The 11 overpayment data points collected by Myers and Stauffer fit quite well to a lognormal distribution.[25]

Calculating the lower bound of a one-sided 90% confidence interval for the C Care data while assuming a lognormal distribution can be done in several defensible ways. They all result in a lower bound that happens to be similar to the lower bound Myers and Stauffer calculated. In other words, assuming a lognormal distribution puts the statistical analysis on firmer footing, but it does not change the result in this particular case.[26] Although a lognormal distribution with a large sample might produce a lower bound closer to the sample mean (potentially yielding a result much more unfavorable to C Care), the small number of data points for overpayment size (11) causes a relatively wide confidence interval.

An entirely different way of approaching the Myers and Stauffer sample is to view it as a single distribution—96 data points—but to convert the overpayments from raw numbers to percentages of the underlying claims from which they were derived. Thus, for example, claim

---

[21]   This was the nature of the Myers and Stauffer audit. The auditor set out to "Verify from documentation that services were actually provided; verify dates of service, procedure codes, revenue codes, drug codes, site of service codes, diagnosis codes; verify medical necessity and/or authorization of services." A.R. 10.
[22]   Direct written testimony of Eschenbach at 2; direct written testimony of Lehman at 8.
[23]   Direct written testimony of Eschenbach at 4.
[24]   Oral testimony of Eschenbach. In this kind of distribution, the logarithm of the variable is normally distributed, even though the variable itself is not. C Care Exhibit 4 gives examples of lognormally-distributed data.
[25]   Direct written testimony of Eschenbach at 2.
[26]   *Id.* at 3, 6-7.

D908709, the one with a $409.50 overpayment, becomes a data point of 60.5% because it represented an overcharge of 60.5% on a $677.25 claim. The overall methodology is to generate a 90% one-sided confidence interval for what the true mean *percentage* of overpayment is across the universe of claims. In this context D908709 is not an outlier, but rather is only the third highest data point, and no single data point accounts for more than about 20% of the mean overpayment calculated from the sample.

With a large number of 0% data points, the sample remains skewed, but its skewness[27] and kurtosis[28] fall to much lower levels. This method has some intuitive attraction, in that the result of the sample does not depend as heavily on any one data point. The method would nominally be applied using a normal distribution assumption but with recognition that the data fits no particular distribution, and in this case it would result in a mean overpayment percentage of 4.94%[29] and a 90% lower confidence limit of 2.62%.[30] If accepted, this calculation would be very unfavorable to C Care, in that 2.62% of C Care's 2007 claims translates to about $104,000, much more than the $45,683.20 that Myers and Stauffer thought to be the 90% lower confidence limit.[31]

The Division's expert ultimately recommended this statistical approach as the one the Division *ought* to be using (though he did not opine that it is the only valid approach). The trouble with applying it in this case, however, is that the statistical text deemed authoritative by both Drs. Lehman and Eschenbach makes it quite clear that the sample size needed to be confident of the lower bound, even with the lower kurtosis and skewness shown when this sample is converted to percentages, is about 189 claims, not 96 claims.[32]

The usefulness of looking at the data through this prism is not entirely destroyed, however, by the fact that the sample size is indisputably too small to meet textbook standards. The trend beginning to appear after half the necessary sample has been drawn is a trend that confirms, rather than contradicts, the conclusion drawn from other methods of analyzing the sample. One way to see this trend is to suppose that Myers and Stauffer had indeed drawn a sample of 189 claims, but all of the 93 claims to be hypothetically added to the sample were zero

---

[27]   "Skewness" is a term used in statistical theory to refer to "how much a distribution of values departs from symmetry around the mean." Direct written testimony of Lehman at 3.
[28]   "Kurtosis" is a measure of the prevalence of outliers. *Id.*
[29]   474.7 (the total of the overpayment percentages) ÷ 96.
[30]   September 11, 2011 written testimony of Eschenbach.
[31]   *Id.*
[32]   *Id.*; September 26, 2011 written rebuttal testimony of Lehman.

percent overpayments, simulating the most favorable possible outcome to C Care of a larger sample size. The result would be a mean overpayment percentage in the sample of 2.51%,[33] which translates to a mean overpayment total that would still be over $100,000.[34]

In short, the method by which Myers & Stauffer calculated the lower bound of a one-sided 90% confidence interval was not statistically rigorous. However, the more sophisticated analysis offered by the Division's expert tends to confirm the overall Myers and Stauffer result. In fact, the Myers and Stauffer method "is likely to under-estimate the lower bound on the average cost,"[35] resulting in a more favorable result for C Care than would have been generated by a better statistical model.

The Myers and Stauffer sample of 96 claims was drawn randomly. It is sufficient, when valid statistical techniques are applied, to support a lower bound for the one-sided 90% confidence interval approximately equal to or greater than the $45,683.20 lower bound that the auditor identified. It firmly establishes that the full universe of C Care's 2007 claims, more likely than not, contains overcharges totaling more than $45,683.20.

IV.   Analysis

The hearing offered and provided in this case was a *de novo* hearing, in which the Division undertook to prove the amount owed. Such a hearing is designed to reach, for the first time, a final, commissioner-level decision on whether recoupment is due and, if so, in what amount.

What the Division had to prove to prevail on its claim was that it is more likely than not that C Care received overpayments of $45,683.20 or more. Nothing in statute or regulation requires the Division to prove the recoupment amount to a 90% probability.

By regulation, the Division is permitted to make its case by means of statistically valid sampling methods. It is the Division's *policy* to seek only to recoup the amount at the lower

---

[33] $474.7 \div 189$.

[34]   Of course, this hypothetical sample would have a larger proportion of zero values and therefore slightly greater skewness and substantially greater kurtosis than the 96-claim sample actually taken, and so the sample size would still be inadequate by textbook standards. Accordingly, no effort is made here to approximate what the 90% one-sided lower confidence interval would be. *Cf.* September 11, 2011 written testimony of Eschenbach (Stine & Foster guidelines would require a sample size 10 times the skewness squared and 10 times the absolute value of the kurtosis).

[35]   Direct written testimony of Eschenbach at 4. This is because the Myers and Stauffer method assumes normal distribution with a left-hand tail similar to the right-hand tail, whereas the probable distribution is lognormal and asymmetrical and would "pull" the lower bound to the right. However, this is not quite the same as saying the Myers and Stauffer method is "conservative" in every case. It is simply the wrong model, and the Division may wish to consider shifting to one of the more apt models identified during this case.

bound of the 90% one-sided confidence interval that can be validly computed from the sample. It was because of this policy that the evidence in this case focused on that lower bound.

Through the hearing process, the Division proved that a representative sample of 96 claims has been validly drawn at random and validly analyzed for billing errors. Although 96 claims is a relatively small sample, it is certainly not the smallest that has been used and upheld by courts in audits of this kind.[36] The Division also proved, by means of valid statistical analysis—albeit not the analysis done by Myers and Stauffer—that the sample is sufficient to make it more likely than not that C Care owes at least $45,683.20.

Two factors bolster this conclusion. The first is the conservatism built into the Division's methodology. By striving for 90% confidence, the Division makes allowance for the imperfection of matching a statistical model to a particular real-world situation. Further, the use of such an interval imposes an ever greater discount on the Division's claim, the smaller the sample used. In this case, it reduced the Division's claim from the sample mean of $9.35 per claim, or more than $131,000, to barely one-third of those amounts: 3.26 per claim and $45,683.20.

The second factor that bolsters this conclusion is the opportunity for a full, *de novo* hearing. At least as the Division has applied the program in this case, a provider who feels statistical analysis has produced a misleading result has the option to have an auditing firm draw and analyze a different or expanded sample, or to provide a more credible interpretation of the sample already drawn. Where that does not occur (or, more precisely, where the results of a competing sample are not presented), it may be reasonable to infer that the provider prefers to rest on the conservative analysis of the existing sample rather than arrive at the true rate of error, which may be considerably less favorable. Courts in other jurisdictions have found that an opportunity for a provider to rebut the statistically-derived conclusion is a key element in making such a procedure fair.[37] The opportunity provided by a full *de novo* hearing is as extensive as any that can be offered.

---

[36] *See, e.g., Maxxim Care EMS, Inc. v. Sebelius,* 2011 WL 5977666 (S.D. Texas) (30-claim sample); *Michigan Dep't of Educ., supra,* 875 F.2d at 1199 (overall sample of 0.4% of population—smaller in relative terms than the 0.7% sample taken from C Care claims; individual strata sampled at approximately 50 claims per stratum).
[37] *See, e.g., Ratanasen v. California Dep't of Health Services,* 11 F.3d 1467, 1471-72 (9th Cir. 1993); In re Bailey, 581 N.E.2d 577, 580 (Ohio App. 1989).

V.   Conclusion

The Division has shown that, more likely than not, C Care Services LLC submitted at least $48,612.70 in overbillings in 2007, comprised of the $45,683.20 discussed in Parts III and IV above and the additional uncontested amounts shown in the table on page 2. Accounting for reimbursements already made, the Department of Health and Social Services may recoup $39,961.59 from C Care.

DATED this 10th day of April, 2012.

By: *Signed*_____
         Christopher Kennedy
         Administrative Law Judge

Adoption

The undersigned adopts this decision as final under the authority of AS 44.64.060(e)(1). Judicial review of this decision may be obtained by filing an appeal in the Alaska Superior Court in accordance with AS 44.62.560 and Alaska R. App. P. 602(a)(2) within 30 days after the date of this decision.

DATED this 15th day of May, 2012.

By: *Signed*_____
         Signature
         William J. Streur_____
         Name
         Commissioner_____
         Title

[This document has been modified to conform to the technical standards for publication.]