Uncertainty Analysis of Analytical Results When Errors are Not Normally Distributed*

Allen Brodsky**, Daniel G. Martinez, and W. Jeffrey Klemm
Science Applications International Corporation
1410 Spring Hill Road, Mail Stop SH2-1
McLean, VA 22102

Abstract

This paper presents a method for calculating *a priori* uncertainty intervals for plutonium interpretations by fission track analysis (FTA) of human urine samples. The derivations of the method are adaptable to estimating uncertainty intervals of analysis of single samples, based on calibration data obtained from multiple samples spiked over a range of concentration. The data used in last year's BAER conference talk on the population excretion of plutonium had approximately normally-distributed errors in tracks observed at given spike levels, with almost constant variance in tracks vs. plutonium spike level. This allowed fitting of the pre-September 1998 calibration line by a simple linear regression. Methods are presented here for estimating *a priori* uncertainty intervals about interpretations using post-September 1998 data, which had non-normal error distributions and variance increasing with spike level.

INTRODUCTION

At last year's meeting, a mathematical characterization of the reported distribution of plutonium in one-day urine samples from a population exposed to plutonium only from global fallout was reported (Brodsky 2000). The use of the lognormal fit to the upper part of the cumulative distribution of this data has been used as one source for obtaining an estimate of the population dose from plutonium in fallout (Barss et al. 2001). When that FTA data were obtained (pre-September 1998), the calibration data showed approximately normal distributions of error and constant variance versus plutonium spike level (Schaeffer et al. 1999). After September 1998, the precision of the FTA analysis deteriorated significantly (due to laboratory changes beyond the chemist's control), and the error distribution was no longer normal, nor was the variance constant with plutonium level. Thus, a simple linear regression (least sum-of-squared deviations) fit would not be a "best fit" characterization of the calibration data, in the sense that a goodness-of-fit test would not yield a minimum reduced chi-square statistic. Klemm et al. (2001) have described the derivation of a best fit calibration line to the post-September 1998 data, based on the finding that the ratios of observed tracks divided by tracks expected from the best fit line were lognormally distributed.

This present paper derives uncertainty distributions of interpretation at each level of plutonium in urine, using sample calibration data with non-normal error distributions as fitted by Klemm et al. (2001). The uncertainty intervals are compared with analogous "discrimination intervals" derived in Mood (1950, pp.299-301) under the assumption of normally-distributed errors.

METHODS AND RESULTS

Most of the defined symbols and equations used in derivations and calculations in this paper

were prepared using MATHCAD™. (MATHCAD is a trademark of MathSoft, Inc.) This format has the advantages of reducing the complexity and length of the presentation, and providing the mathematical algorithms in an exact form that can be easily read and adapted by others. For simplicity, the parameter definitions and formulations for the lognormal distributions used in this paper are obtained from Aitchison and Brown (1963), Gilbert (1987), or Brodsky (1982). Caution is necessary in reading symbols: 1) The symbol s_g is used for the sample "standard geometric deviation" (called "geometric standard deviation" in Boecker et al. (1991)) – this is the ratio of the 84.13 percentile to the 50 percentile, the ratio of the 50 percentile to the 15.87 percentile, or the average of both; 2) The symbol σ_g is **not** in this paper the underlying "true" population value of the sample estimator s_g ; it is defined as the sample estimate of the standard deviation in the logarithmic-transformed x value; i.e., the relationship is $\sigma_g = \ln s_g$. This avoids excessive use of the "hat" symbol to designate sample estimators.

Figure 1 shows a sample of pre-September 1998 FTA calibration data that was used by Brodsky et al. (1999, 2000a, 2000b) in mathematically characterizing a control population urinary excretion distribution, and Barss et al. (2001) for obtaining, together with autopsy data, estimates of the U. S. population internal dose from fallout plutonium. Recognizing the unfortunate changes that had occurred in laboratory conditions, which must be extremely stringent and quality-controlled for best FTA results (Boecker et al. 1991), the post-September data were plotted separately as in Figure 2. Figure 1 shows an approximately constant variance with plutonium level (homoscedasticity), a small enough variability so that normal error distribution assumptions are reasonable, and an intercept at the Y axis close to the true mean value of the simulated urine blank distribution. However, in Figure 2, homoscedasticity is lost, and the intercept with the blank data included still does not agree with the mean blank value. Thus, it seemed that something different must have occurred in the chemistry and/or irradiation procedures for the blanks as compared with the chemical and/or isotopic matrix of the spiked solutions, so only the spiked data are used for the regression and error analysis here. Thus, neither a simple nor a weighted regression, under the usual textbook (Chapman 2000; Miller and Miller 1993; Brodsky 1982) assumptions of normally-distributed errors are applicable.

Figure 3 shows the cumulative plot of y/Y that provided the best model of the error distribution; it is plotted on a probability vs. a natural log scale, by MINITAB algorithms. This indicates that y/Y is lognormally distributed, and thus $\ln (y/Y) = \ln y - \ln Y$ is normally distributed. This was checked by a plot of the deviations in logarithms. Exhibit 1 presents the solution of the equations obtained from the partial differentiations, with respect to slope and intercept, respectively, of a chi-square statistic made up of a sum of squared deviations ($\ln y - \ln Y$) divided by respective variances in these log deviations. The lower part of this exhibit shows that the same values, slope $a = 0.679$ and intercept $b = 24.094$, were obtained by the MINIMIZE algorithm (Klemm et al. 2001). Figure 4 compares the calibration lines obtained by simple regression and this logarithmic y/Y regression. Since interpretations of single samples would be made from the inverse relation, $x = y/a - b/a$, the uncertainty distribution of interpretations must be obtained by an algorithm for the probability density of x at any activity level.

Distribution functions are derived here for the case of a random determination of tracks in a human urine sample, paired with the subtraction of a randomly varying blank. (Previously, the characterization of the distribution of plutonium in control population urines showed similar

numerical probabilities for paired- and constant-blank cases (Brodsky et al. 1999; Brodsky 2000a, 2000b).) For illustrating the approach, the slope, a, is assumed constant over the period of interpretations. As shown in the second line of definitions at the top of Exhibit 2, the standard deviation in logarithms (and the standard geometric deviation) for the blanks was assumed to be the same as found for the spike error distribution of Figure 3. The mathematical forms of the probability densities for the two terms in $x = y/a - b/a$, $q(w)$ where $w = y/a$, and $p(u)$ where $u = b/a$, are presented below the definitions at the top of Exhibit 2, respectively. The form of $q(w)$ was obtained by transforming the lognormal distribution of y/Y, modeled using parameters obtained from plots such as in Figure 3, the form of the lognormal density function, and Theorem 2.1, pages 10-11 in Aitchison and Brown (1963). The form for the cumulative distribution F as evaluated for discrete values of activity is shown on the right side of Exhibit 2. Exhibit 3 shows F defined as $G(z)$ to obtain a continuous function, the forms of which are illustrated below the function. (Special algorithms for $z \leq 0$ are not shown.) Exhibit 4 shows an algorithm for finding the upper and lower bounds of the 90% interpretation range for a single human urine sample, as evaluated at the 200 aCi actual level. Figure 5 shows the uncertainty distributions of interpretations for 50, 100, 200, and 300 aCi. Exhibit 5 shows an algorithm for solving the three non-linear equations in Mood (1950), for obtaining the lower bound of the 90% "discrimination" interval at the 200 aCi level if the data had been assumed normally distributed.

CONCLUSIONS

1. Error distributions should be examined under the stable analytical conditions to be used for a set of samples before the method of fitting a calibration curve is selected.
2. For the data set used for illustration here, a simple regression calibration line would give interpretations about 20% to 10% too low over the range 100 to 300 aCi, compared to those obtained by the log y/Y error regression (see Figure 4).
3. The 90% uncertainty range using the log y/Y regression at 200 aCi is (47.1, 549), compared to the Mood (1950) discrimination range of (-53, 461) if normally-distributed errors had been assumed here.

References

Aitchison, J.; Brown, J. A. C. The lognormal distribution. Cambridge, England: Cambridge at the University Press; 1963.

Barss, N. M.; Brodsky, A.; Jackson, E. Estimating population doses from plutonium from fission track analysis and autopsy data. Presented at the annual meeting of the Health Physics Society. Health Phys., Supp., 2001.

Boecker, B.; Hall, R.; Inn, K.; Lawrence, J.; Ziemer, P.; Eisele, G.; Wachholz, B.; Burr, Jr., W. Current status of bioassay procedures to detect and quantify previous exposures to radioactive materials. Health Phys. 60, Sup.1:45-100; 1991.

Borak, T. B., editor. Applications of probability and statistics in health physics. Madison, WI: Medical Physics Publishing; 2000.

Brodsky, A. Statistical methods of data analysis. In: Brodsky, A., editor, Handbook of radiation measurement and protection, Section A, Vol. II. Boca Raton, FL: CRC Press, Inc.; 1982, pp. 261-330.

Brodsky, A.; Schaeffer, D. M.; O'Toole, S.; Kaplan, E.; Barss, N. M.; Dancz, J.; Klemm, W. J.; Raine III, D. A.; and Stiver, J. Statistical model for fission track analysis of plutonium in human samples. Health Phys. 76(6), S175; 1999.

Brodsky, A. Cumulative lognormal distributions of dose-response vs. dose distributions. In the proceedings of the 10th International Congress of the International Radiation Protection Association, Hiroshima, Japan: May 14-19, 2000a.

Brodsky, A. Mathematically analytic distributions can explain radioanalytic results. In the proceedings of the 46th Annual Conference on Bioassay, Analytical and Environmental Radiochemistry. Seattle, WA: November 12-17, 2000b.

Chapman, P. L. Applied regression methods. In: Borak, T. B. *op. cit*.; 2000.

Gilbert, R. O. Statistical methods for environmental pollution monitoring. New York: Van Nostrand Reinhold Company; 1987, pp. 168-169.

Health Physics Society. Performance criteria for radiobioassy. McLean, VA: Health Physics Society, American National Standards Institute Standard ANSI N13.30-1996; 1996.

Klemm, W. J.; Brodsky, A.; Schaeffer, D. M. Radioanalytic data interpretation when the ratio reading/median is lognormally distributed. Presented at the annual meeting of the Health Physics Society. Health Phys., Supp.; 2001.

Miller, J. C.; Miller, J. N. Statistics for analytical chemistry, 3rd edition. New York: Ellis Horwood PTR Prentice Hall; 1993.

Schaeffer, D. M.; O'Toole, S.; Kaplan, E. Calibration and interpretation of fission track analysis of plutonium in urine. Health Phys. 76:S175; 1999.

- For presentation at the 47th Annual Conference on Bioassay, Analytical and Environmental Radiochemistry, Honolulu, Hawaii, November 4-8, 2001. The views expressed in this paper are those of the authors alone. Also, the data presented are not necessarily representative of any appropriate sample used for an official purpose.

** Also, Adjunct Professor, Georgetown University, Washington, DC.