Draft

ADVISORY BOARD ON RADIATION AND WORKER HEALTH

National Institute for Occupational Safety and Health

DRAFT REVIEW OF ORAUT-RPRT-0044: ANALYSIS OF BIOASSAY DATA WITH A SIGNIFICANT FRACTION OF LESS-THAN RESULTS

Contract No. 200-2009-28555
SCA-TR-PR2010-0009, Revision 0

Prepared by

Joyce Lipsztein, PhD
Harry Chmelynski, PhD
S. Cohen & Associates
1608 Spring Hill Road, Suite 400
Vienna, VA 22182

November 2010

Disclaimer

This document is made available in accordance with the unanimous desire of the Advisory Board on Radiation and Worker Health (ABRWH) to maintain all possible openness in its deliberations. However, the ABRWH and its contractor, SC&A, caution the reader that at the time of its release, this report is pre-decisional and has not been reviewed by the Board for factual accuracy or applicability within the requirements of 42 CFR 82. This implies that once reviewed by the ABRWH, the Board's position may differ from the report's conclusions. Thus, the reader should be cautioned that this report is for information only and that premature interpretations regarding its conclusions are unwarranted.

S. COHEN & ASSOCIATES: ***Technical Support for the Advisory Board on Radiation & Worker Health Review of NIOSH Dose Reconstruction Program***	Document No. SCA-TR-PR2010-0009 Effective Date: Draft – November 12, 2010
DRAFT REVIEW OF ORAUT-RPRT-0044: ANALYSIS OF BIOASSAY DATA WITH A SIGNIFICANT FRACTION OF LESS-THAN RESULTS	Page 2 of 21
Task Manager: _____ Date: _____ Stephen F. Marschke	Supersedes: N/A
Project Manager: _____ Date: _____ John Mauro, PhD, CHP	Reviewers: Arjun Makhijani, PhD Stephen F. Marschke

Record of Revisions

Revision Number	Effective Date	Description of Revision
0 (Draft)	11/12/2010	Initial issue

TABLE OF CONTENTS

ABBREVIATIONS AND ACRONYMS

CL	Censoring Level
CW	Construction Worker
dpm	disintegrations per minute
EPA	U.S Environmental Protection Agency
GSD	Geometric Standard Deviation
HLN	hybrid lognormal
L	Liter
L_D	detection limits
MDA	Minimum Detectable Activity
NCW	Non-Construction Worker
NIOSH	National Institute for Occupational Safety and Health
NOCTS	NIOSH OCAS Claims Tracking System
NRC	U.S. Nuclear Regulatory Commission
ORAUT	Oak Ridge Associated Universities Team
OTIB	ORAUT Technical Information Bulletin
PIC	Pocket Ionization Chamber
POC	Probability of Causation
PROC	Procedure
RPRT	Report
SEC	Special Exposure Cohort
SRS	Savannah River Site
TLD	Thermoluminescent Dosimeter

EXECUTIVE SUMMARY

This report presents SC&A's initial review of ORAUT-RPRT-0044, *Analysis of Bioassay Data with a Significant Fraction of Less-Than Results* (ORAUT 2009), which describes the National Institute for Occupational Safety and Health's (NIOSH's) proposed methods for reconstructing exposures to unmonitored workers when essentially all or most[1] of the available data from monitored workers are below the limit of detection. The proposed methods complement and supplement the recommendations contained in ORAUT-PROC-0095, *Generating Summary Statistics for Coworker Bioassay Data* (ORAUT 2006a). This report also addresses potential issues related to the application of the proposed methods to construction workers at the Savannah River Site (SRS), especially those issues that might be considered SEC-related. This report does not contain a detailed analysis of the issues, and is intended solely for use by the designated working groups.

SC&A has identified the following four findings.

Finding No. 1: The statistical methods proposed in ORAUT-RPRT-0044 are based on sound statistical methodologies, and the material is well presented. The proposed methods are an improvement over the regression methods proposed in ORAUT-PROC-0095 when essentially all or most of the data are less-than results, the limit of detection was the same for all samples in the dataset, and **the samples above the limit of detection are randomly spread across workers, job types, and work areas**.

The work location and work assignments of the workers with positive results are not considered in the NIOSH approach. Before these methods are used in a coworker model, further analysis of the positive results is required. In particular, identification of the workers, work areas, and processes accounting for the positive results in the datasets is required to reveal possible patterns that may explain the occurrence of positive results.

NIOSH does not offer any consideration relating to the pattern or time distribution of the positive results. For example, the positive results could be present x times per year, during defined periods of time, or during a specific campaign. It is possible that the same subgroup of workers accounted for most of the positive readings year after year.

Finding No. 2: ORAUT-RPRT-0044 does not address the representativeness of the dataset for workers in all work areas and job types. No individual worker analysis was performed, as the report concentrates only on analysis of a collection of analytical results.

Finding No. 3: The methods proposed in RPRT-0044 are illustrated using several examples. One example uses simulated data from two distributions. In this situation, the maximum likelihood method performs as expected. In a second example involving analysis of SRS Pu-239 urinalyses, detection limits thought to reflect conditions before the 1980s are applied retrospectively to a 1997 dataset that contained no less-than results. The 1997 dataset contains

[1] SC&A interprets "essentially all" as greater than 95% of the results below the censoring level, and "most" as between 90% and 95% of the results below the censoring level.

no less-than results, due to the capability of modern laboratory procedures. The 1997 SRS data selected for the example appear to be a mixture of two different distributions. Although no less-than results are found in the 1997 dataset, the method of RPRT-0044 is illustrated by selecting a hypothetical detection limit higher than the lower of these two distributions. Since the data used in this example do not reflect working conditions at SRS prior to 1980, the use of SRS urinalysis data in this example is largely an academic exercise, providing no useful information about pre-1980 exposures at SRS or the appropriateness of using the two-distribution model for any type of worker in time periods before 1997.

Finding No. 4: The methods proposed in RPRT-0044 for datasets with essentially all or most of the less-than results are based on samples obtained from all workers, regardless of job type or location. No attempt was made to determine the work areas, processes, or job types of workers with positive results. This approach is not claimant favorable for construction workers for three reasons:

(1) In many cases, construction workers were not monitored as frequently as non-construction workers,[2] hence the dataset may not be representative of the distribution of construction worker exposures.

(2) Because constructions workers were sampled less frequently, a higher percentage of these workers will require use of the coworker model.

(3) The positive samples may come from very few workers or restricted time periods, which may not be representative of the worker population.[3]

The work assignments of the workers with samples in the upper tail of the mixture distribution may have an unexpectedly high number of construction workers when compared with their degree of monitoring. The work assignments of workers with samples in the two populations should be inspected and categorized by job type to look for such disparities.

[2] For instance, in the data complied by NIOSH for addressing SRS CW vs. NCW plutonium exposures for ORAUT-OTIB-0052, there were 612 NCW (roll numbers 1 to 3) and 888 CW (roll numbers 4 to 6), yet only about 30% of the samples were from CWs. Moreover, there were no CW samples in the 1950s and few in the 1960s and 1970s. Most of the samples were in the 1980s. In contrast, the NCW samples were spread throughout the period from 1953 to the mid-1990s (see SC&A 2010, Figure 1). The numbers of CWs and NCWs were determined from the NIOSH spreadsheet where the data were compiled. See also Figure 3 below, which shows that CWs at SRS were less frequently monitored for plutonium in most of the years prior to 1990.

[3] Both were true of the CW plutonium bioassay dataset used by NIOSH for its SRS analysis in ORAUT-OTIB-0052 (see SC&A 2010, pp. 5 to 8).

1.0 INTRODUCTION

This report presents methods for analyzing datasets dominated by censored results. Methods are presented for two types of datasets; (1) those that have a large fraction (from 90% to 95%) of censored data, and (2) datasets that consist entirely (>95%) of censored data. Censored data is a statistical term for measurements below the limit of detection, i.e., the "Less-Than Results" noted in the title of the report. This report is intended to correct/complement ORAUT-PROC-0095, *Generating Summary Statistics for Coworker Bioassay Data* (ORAUT 2006a). The methods recommended in PROC-0095 for estimating parameters of a lognormal distribution are applicable when less-than results are present in the data, but there must be a sufficient number of measurements above the detection limit to fit a regression line to the data.

The PROC-0095 approach assumes that there is a single lognormal distribution applicable to the entire dataset that is being analyzed. In the case where there are no values above the detection limit, PROC-0095 cannot be applied. If there are only a few values above the detection limit, the regression method in PROC-0095 may yield erratic results. If the dataset contains more than one population with possibly different distributions, PROC-0095 is no longer applicable. RPRT-0044 addresses these situations where the procedure recommended in PROC-0095 is not applicable, cannot be applied, or may yield erratic results.

The deficiencies of ORAUT-PROC-0095 when there is a large number of less-than results were noted by SC&A in Finding ORAUT-PROC-0095-03 in *Findings from 3rd set of Procedures*. In *NIOSH Responses to Selected Findings from 3rd set of Procedures*, NIOSH explains that a new report (ORAUT-RPRT-0044) has been developed to better address the issue of censored data. The statistical treatment of the data in RPRT-0044 is correct and well explained. If the dataset contains urine results for which most samples do not have analyte in the urine, but a small fraction do, the methods presented in RPRT-0044 are an improvement over the PROC-0095 procedure.

RPRT-0044 consists of two separate methods, to be applied depending on the relative fraction of less-than results in the dataset. The first method is described in Section 2 of the report. It relates to datasets consisting of results that are essentially all below the censoring level (CL). The statistical treatment of the data is correct and well explained. The general approach is to determine a normal distribution to represent the measurements that lie below the CL. The upper tail of this normal distribution is then approximated by a lognormal distribution for use in dose modeling.

The second method presented in RPRT-0044 (Section 3 of the report) addresses the case when the number of results above the CL is larger, but a large majority of results are less than the CL. The approach in this case is to assume that there are two distinct populations; one group of samples has no measureable analyte in the urine, but a small fraction of the samples do. The samples in the former group are assigned a normal distribution of "background" readings, and the samples in the latter group are assigned a lognormal distribution. Maximum likelihood methods are used to estimate the parameters of the mixed model.

Section 4 of RPRT-0044 contains an example based on the use of 1997 SRS Pu-239 urinalysis data. The use of this dataset to illustrate the proposed methods is largely an academic example, because the dataset actually contains no less-than results, due to the advanced capabilities of modern analytical techniques and advances in process technology and worker protection. It is likely that the few positive results in this dataset are incident-related.

The "two-population" method proposed in RPRT-0044 is illustrated by assuming a high CL thought to be reflective of analytical techniques used in the years before the 1980s. Since the data used in this example are 1997 data, it is unlikely that the data reflect working conditions at SRS in the earlier years. The two-distribution approach may only be useful in the analysis of the most recent datasets. RPRT-0044 provides no evidence of the existence of a "two-population" model in datasets from earlier years.

2.0 REVIEW OF SECTION 2 OF ORAUT-RPRT-0044

This section of RPRT-0044 presents a method that can be applied to datasets consisting of results that are essentially all below the CL. The general approach is to determine a normal distribution to represent the measurements that lie below the CL. The upper tail of this normal distribution is then approximated by a lognormal distribution for use in dose modeling.

2.1 GENERAL COMMENTS

In the application of the procedure recommended in RPRT-0044, the issue of completeness of the available data has not been addressed. Some workers were not monitored, otherwise there would be no need for a coworker model. The underlying assumption seems to be that the workers with the most exposure potential were monitored; but we have seen in a number of cases that this was not necessarily true. So, if essentially all known results are below the MDA, then can the whole population (including the unmonitored workers) be represented by a distribution that is almost entirely below MDA, except for some scattered people in a lognormal tail fit to the same distribution? If the unmonitored workers are from a different population, the applicability of a coworker model derived from monitored workers would be in question.

Another limitation of the approach recommended in RPRT-0044 is that it assumes a single detection limit for all measurements in the dataset. Hence, the method is best applied when all measurements were made using the same type of equipment and analytical techniques for the same radionuclides. The method described and exemplified in the report uses only 1 year of results. While the use of 1-year results is consistent with the NIOSH coworker modeling effort, which generates dose estimates for each year, in most cases, the limit of detection of the equipment for the specific analyte and solubility type in question is not known and must be assumed or inferred from other sources.

Many NIOSH datasets contain results noted as “less-than” with no detection limit specified. Often there is no indication of a “less-than” result; only an entry of 0, which is usually interpreted as a less-than result. In these and other cases, NIOSH has determined detection limits retrospectively. For example, the following detection limits are recommended for use at Y-12 (ORAUT 2006b):

> *5A.3 IN VITRO DETECTION LIMITS*
> *Tables 5A-1 and 5A-2 summarize information developed in Section 5.2. The tabulated values for urinalysis results represent laboratory detection limits L_D and do not include uncertainties introduced by sample collection or conversion from submitted volumes to daily void volumes. As noted in Section 5.2, L_D values for some historical techniques remain to be identified, and will be reported in subsequent revisions as available.*

Table 5A-1. Uranium urinalysis detection limits.

Method	*Period*	*Detection limit (mass)*	*Detection limit (activity)*
Fluorometry	*1950–9/1989*	*7 µg/d*	*11 dpm/d*
Gross alpha	*1950–1964*		*47 dpm/d*
Gross alpha	*1965–9/1989*		*26 dpm/d*
Alpha spectrometry	*10/1989–present*		*0.15 dpm/d*

Table 5A-2. Other in vitro detection limits.

Analyte/solubility type	*Method*	*Period*	*Detection limit (activity)*
Tritium (HTO)	*Liquid scintillation*	*10/1988–present*	*2,000 dpm/d*[a]
Isotopic plutonium/M, S	*Alpha spectrometry*	*10/1989–present*	*0.025 dpm/sample*
Am-241/M	*Alpha spectrometry*	*10/1989–present*	*0.050 dpm/sample*
Th-228/M, S	*Alpha spectrometry*	*10/1989–present*	*0.150 dpm/sample*
Th-232/M, S	*Alpha spectrometry*	*10/1989–present*	*0.070 dpm/sample*
Np-237/M	*Alpha spectrometry*	*10/1989–present*	*0.100 dpm/sample*

a. Estimate

At SRS, detection levels are described in terms of *reporting levels*. The reporting level is described as follows:

> *The reporting level is distinguished from the MDA. As described in WSRC-IM-90-139* [WSRC 1990], *the reporting level is the minimally acceptable decision level; the bioassay laboratory was required to achieve decision levels below the reporting level. SRS technical documentation provided no quantitative relationship between the reporting level and MDA (p. 65, ORAUT 2005).*

Recommendations for reporting levels for plutonium and americium at SRS are shown in the following table, extracted from Appendix D of ORAUT-TKBS-0003 (ORAUT 2005):

Table D-1. Limits of detection for urinalysis

Radionuclide	*Period*	*MDA (pCi/L)*	*Reporting Level*
$^{238, 239, 240}$Pu	*1954 to 1962*		*0.05dpm/1.5L*
$^{238, 239, 240}$Pu	*1963 to 1981*		*0.1dpm/1.5 L*
^{238}Pu	*1981 to 1988*[d]		*0.05dpm/1.5L*
^{238}Pu	*1988 to present*		*0.07 dpm*
^{239}Pu, ^{240}Pu	*1981 to 1988*[d]		*0.07 dpm/1.5 L*
^{239}Pu, ^{240}Pu	*1988 to present*		*0.06 dpm*
^{241}Am	*Mid 1960s to 1971*		*3 dpm/1.5 L*
^{241}Am	*1971 to 1990*		*0.3 dpm/1.5 L*
^{241}Am	*1990 to 1994*		*0.1 dpm/L*
^{241}Am	*1994 to present*	*0.029*	

NOTE: Only the Pu and Am data in Table D-1 are reproduced here.

When the procedure recommended in RPRT-0044 is applied to samples collected in a single year and analyzed using the same equipment, the disparity in limits of detection across samples may be reduced, but a large degree of uncertainty remains in the assignment of a specific value for the limit in that year. In most cases, the limit of detection of the equipment for the specific analyte and solubility type in question is not known, and must be assumed or inferred from other sources. It is unlikely that the detection limits remained constant for long time periods of several decades, as suggested by the tables above. When the dataset does have limits of detection associated with the less-than results, these limits may have been added to the dataset at a later date during a retrospective analysis, such as the one above. Hence, the mere fact that the detection limits in a dataset are all the same for a given radionuclide may indicate only that they were imposed by assumption.

In the early years, with higher limits of detection, there may also be a relatively smaller proportion of measurements above the limit. Lack of detailed information on the type and limit of detection of the equipment in use in these earlier years serves to compound this problem.

2.2 TECHNICAL COMMENTS

A normal distribution is symmetric, while a lognormal distribution is asymmetric and has a longer tail on the right. For sample sizes that are not too large, the samples that fall in the upper tail of a normal distribution may appear to be lognormally distributed. Figure 1 shows a case with N = 1,000 samples from a standard normal distribution with n = 34 of the samples occurring in the upper 5% tail of the distribution (× >1.645 standard deviations). The vertical scale is logarithmic and the regression line fits relatively well, indicating that the samples in the tail of the normal distribution appear to be a sample from a censored lognormal distribution.

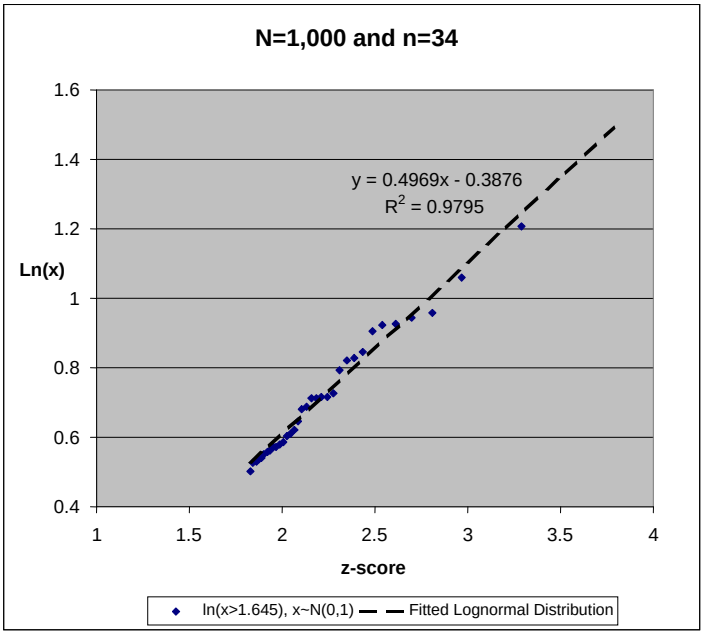

Figure 1: Fitting a Lognormal Distribution to 34 Samples in the Upper Tail of 1,000 Samples from a Standard Normal Distribution N(0,1)

It is well known that the lognormal distribution has "fatter" tails than the normal distribution, but at a moderate sample size of 1,000, the two distributions appear very similar. When the sample size is increased to N = 10,000 with n = 468, then the difference in the upper tails of the two distributions becomes apparent, as shown in Figure 2.

The slope of the fitted regression lines in Figures 1 and 2 are $\sigma = 0.4969$ and $\sigma = 0.4494$, respectively. The geometric standard deviations (GSD = e^{σ}) of the two lognormal distributions are 1.64 and 1.57, respectively. When the upper 5% tail of a sample from the standard normal distribution is fit to lognormal distribution, the expected GSD is 1.55 ($\sigma = 0.438$). This curious relationship between the normal and lognormal distributions forms the gist of the approach recommended for the case of essentially all less-than results: use a lognormal distribution with a GSD of 1.55 and then to match the CL.

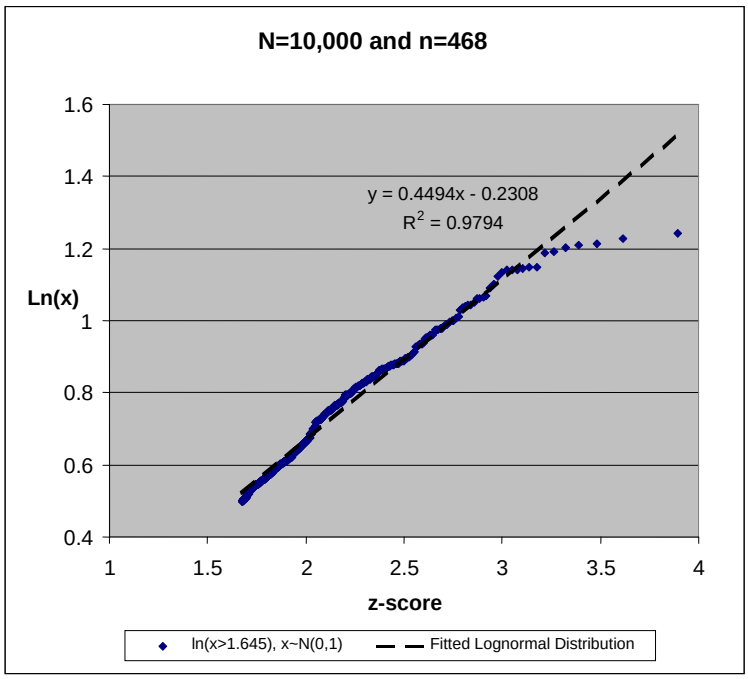

Figure 2: Fitting a Lognormal Distribution to 468 Samples in the Upper Tail of 10,000 Samples from a Standard Normal Distribution N(0,1)

2.3 ISSUES OF IMPORTANCE TO SRS CONSTRUCTION WORKERS

The methods proposed in RPRT-0044 for analyzing datasets that consist almost entirely of less-than results are statistically sound, but use of this approach in the coworker model requires additional qualification that the dataset is representative of all groups of workers. The job types of the workers in the SRS 1997 Pu dataset were not analyzed in RPRT-0044. In SC&A's study of construction versus non-construction worker claimants in NOCTS, the relative frequency of Pu urinalysis samples were compared for these two groups of claimants. Job types are known for most NOCTS claimants.

The time series plot in Figure 3 shows the percentage of claimants that were sampled by year for these two groups of workers and for all workers combined. In this plot, the percentage of CW claimants with at least one sample was lower than for NCW claimants up to approximately 1990. The lower frequency of CW monitoring is particularly marked in some periods, such as the 1950s and early 1960s.

Figure 4 shows a plot of the average number of samples per year for these two groups of workers. In this figure, the number of samples for each NCW exceeded the average number of samples from CWs, except from 1980 to 1985. In some years, NCWs were over-sampled by a margin of over 2 to 1. If the sampling frequency in the general population of SRS workers was the same as for NOCTS claimants analyzed by SC&A, then the dataset used for analysis of SRS Pu urinalyses is not representative of SRS CWs, due to the higher sampling frequency for NCWs. Note that in both figures and for almost all years, the plot for all workers combined is approximately equal to the NCW plot. This fact again shows that a coworker model based on datasets obtained from NCWs and CWs combined would not be representative of CWs.

In the case of the SRS CWs who were monitored less frequently than NCWs, it is likely that this group of workers forms a separate exposure population. If this were the case, the applicability of a coworker model derived from NCW data would be in question. The problem for the CWs is two-fold; (1) The use of a coworker model would be required for a larger percentage of these workers than for the NCW, and (2) the coworker model would be based on data obtained with an under-sampling of construction workers.

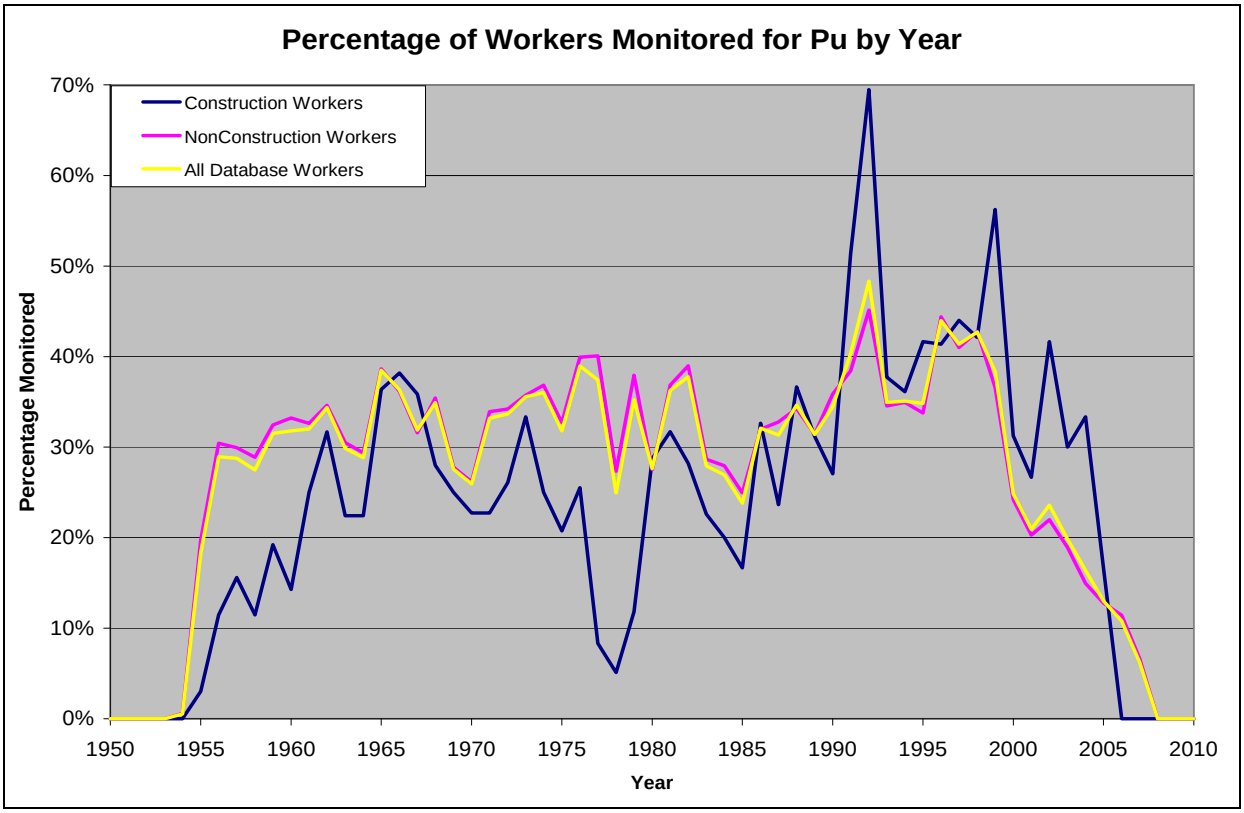

Figure 3: Percentage of Workers Monitored for Plutonium at SRS

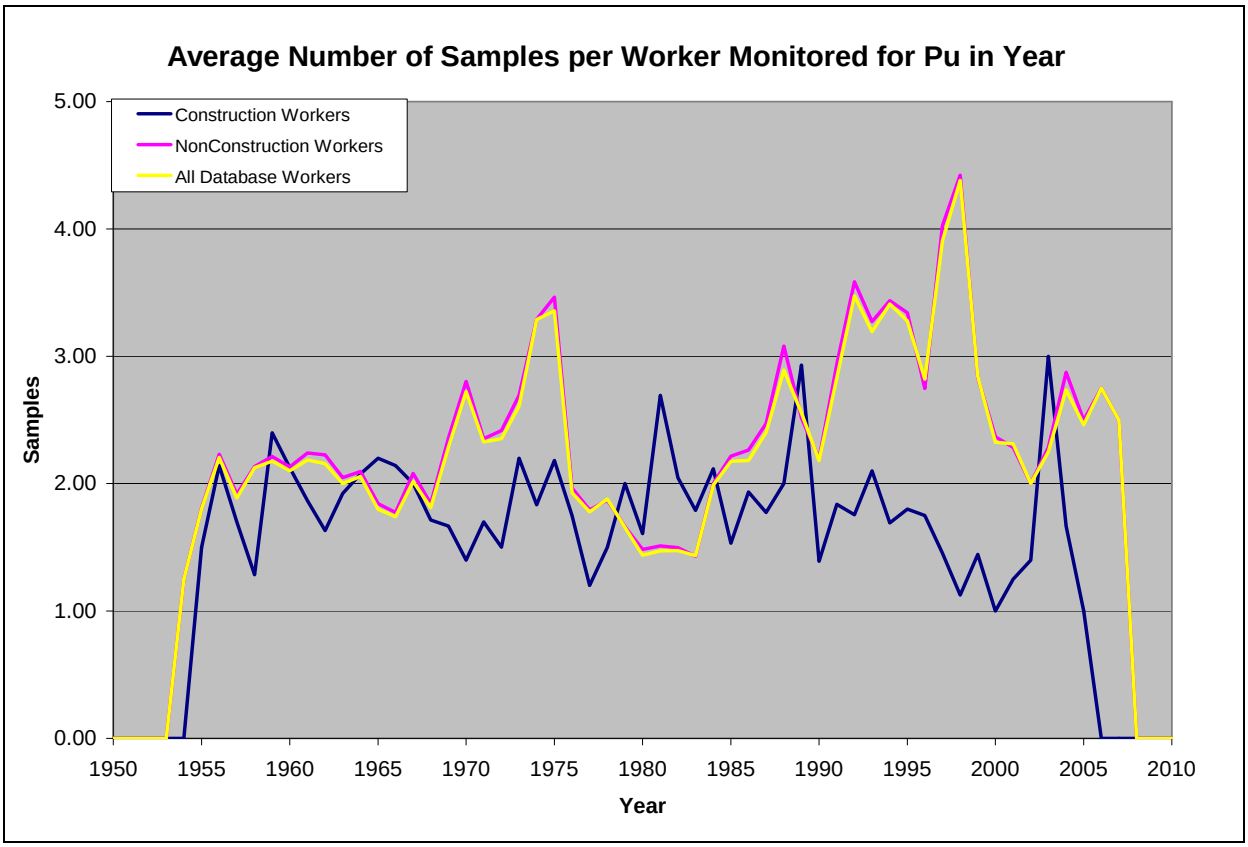

Figure 4: Average Number of Samples per Worker Monitored for Plutonium at SRS

3.0 REVIEW OF SECTION 3 OF ORAUT-RPRT-0044

The second part of the report addresses datasets consisting of results that are mostly below the CL. The authors of the report postulate that the distribution of results is not a single lognormal distribution, but a mixture of two distinct distributions, one normal and the other lognormal. In this model, there are two populations of samples; most samples have no measureable level of analyte in the urine, but a small fraction of the samples do. In this mixed model, the former group of samples with less-than results is assigned a normal distribution representing "background" exposures, and the latter group of samples is assigned a lognormal distribution of exposures.[4]

The statistical treatment of the data is correct and well explained. Maximum likelihood techniques are used to estimate the parameters of the mixed model. Since the method recommended in ORAUT-PROC-0095 for parameter estimation is based on the assumption of a single lognormal distribution, the method recommended there is no longer applicable in this situation.

If a dataset contains urine results for which most of the workers do not have analyte in their urine but a small fraction of the workers do, then the method presented in RPRT-0044 is an improvement over the PROC-0095. Although there are other procedures that may be applied to these datasets, like the Bayesian approach described by Miller et al. (2003), maximum likelihood methods are widely used and are acceptable. This is a valid method to be applied for coworker statistics **when the analytical results above the CL are randomly spread across workers, job types, and work areas**.

The method described and exemplified in the report uses only 1-year results. While the use of 1-year results is consistent with the NIOSH coworker modeling effort, which generates dose estimates for each year, NIOSH does not offer any consideration relating to the pattern or time distribution of the positive results. It is necessary to know if the positive results occur every year, and if those results are related to a particular procedure. For example, the positive results could be present x times per year, during defined periods of time, or during a specific campaign.

The work location and assignments of the workers with positive results are not considered in the NIOSH approach. The method described in the report also does not consider the fraction of the worker population that had positive results. For example, did one sub-group of workers account for most of the positive results? NIOSH does not analyze all the years in which workers were exposed in the facility, where the exposures occurred, or the processes being carried out in the facility. Information of this type is useful to distinguish rare events from ones that appear routinely every year in one part of the facility.

[4] Mixtures of normal and lognormal distributions have been applied by other authors to characterize annual worker exposure. One example is the hybrid lognormal (HLN) distribution proposed by Kumazawa and Numakunai (1981). This model was used by the EPA Office of Radiation Programs to model individual and collective occupational exposures to U.S. radiation workers. In that application, the normal distribution component was used to model exposures in the extreme upper tail, near and above NRC individual exposure limits, while the lognormal distribution component characterized the less-exposed group of workers.

3.1 COMMENTS SPECIFIC TO SRS

The issues of work location and assignment are fundamental in the use of coworker models for CWs at SRS. The work assignments of the workers in the upper tail of the mixture distribution may have an unexpectedly high number of CWs when compared with their degree of monitoring. The work assignments of workers in the two populations should be inspected and categorized by job type to look for such disparities.

Section 4 of RPRT-0044 contains an example that uses 1997 SRS Pu-239 urinalysis data. The use of this dataset to illustrate the proposed methods is largely an academic example, because the dataset actually contains no less-than results, due to the advanced capabilities of modern analytical techniques and recent advances in process technology and worker protection. It is likely that the few positive results in this dataset are incident related.

Plutonium-239 data collected in 1997 at the SRS is used as an example of the "two-population" model. None of this dataset was censored data, using 1997 analytical procedures. When the cumulative distribution of the data is plotted using the graphical technique described in PROC-0095, there appear to be two populations of results. The method of RPRT-0044 is applied retrospectively, by postulating a CL of 0.1 dpm/L and imposing this CL on data collected much later in 1997. This CL is proposed as being representative of analytical procedures for plutonium before the 1980s. This value exceeds the reporting level for Pu-239 of 0.1 dpm/1.5 liters recommended by NIOSH in Table D-1 of ORAUT 2005, reproduced in Section 2.1 of this report.

The example selected has little relation to actual exposures that may have occurred at SRS before the 1980s. In 1997, the data reflect different conditions and processes and many advances in worker safety than before the 1980s. In the 1997 dataset, the relatively few high exposures may have resulted from unusual incidents, while many workers experienced much safer conditions. The clear success of the two-population model for 1997 data does not imply that the same shape of distribution is applicable to SRS before the 1980s.

The two-population model for the 1997 SRS data is described as follows in RPRT-0044:

> *The lower-level distribution represents the normally distributed analytical background. … The higher-level data above the CL of 0.1 dpm/L represent a second population of data in addition to the analytical background. This higher-level population is referred to here as the "exposure population" because it represents results that are unlikely to have come from fluctuations in the analytical background.*

There are several ways to interpret the terms "population" and "analytical background" as they are used in RPRT-0044. The "population" is not a distinct population of workers, but a population of sample results. An individual worker may have samples in both populations.

As noted in Section 2.3 of this report, the dataset used in this example may not be representative of CW exposures at SRS. The difference in the number of samples per worker noted in that

section apply equally to the two-population model. Figure 4 showed a plot of the average number of samples per year for these two groups of workers. In this figure, the number of samples for each NCW exceeded the average number of samples from CWs in 1997 by a margin of over 2 to 1. If the sampling frequency in the general population of SRS workers was the same as for NOCTS claimants analyzed by SC&A, then the dataset used for analysis of SRS Pu urinalyses is not representative of SRS CWs, due to the higher sampling frequency for NCWs.

4.0 REVIEW CHECKLIST

Table 1. Procedure Review Outline/Checklist

Document No.: ORAUT-RPRT-0044	**Effective Date:** 05/22/2009
Document Title: Analysis of Bioassay Data with a Significant Fraction of Less-Than Results	
Reviewer: J. Lipsztein/H. Chmelynski	

No.	Description of Objective	Rating 1-5*	Comments
1.0	**Determine the degree to which the procedure supports a process that is expeditious and timely for dose reconstruction.**		
1.1	Is the procedure written in a style that is clear and unambiguous?	5	
1.2	Is the procedure written in a manner that presents the data in a logical sequence?	3	Only one year of data is analyzed for SRS.
1.3	Is the procedure complete in terms of required data?	1	See comment 1.2.
1.4	Is the procedure consistent with all other procedures that are part of the hierarchy of procedures employed by NIOSH for dose reconstruction?	N/A	
1.5	Is the procedure sufficiently prescriptive in order to minimize the need for subjective decisions and data interpretation?	3	Variation of MDAs over time is not addressed.
2.0	**Determine whether the procedure provides adequate guidance to be efficient in instances where a more detailed approach to dose reconstruction would not affect the outcome.**		
2.1	Does the procedure provide adequate guidance for identifying a potentially high probability of causation as part of an initial dose evaluation of a claim?	2	Several critical issues, such as job types and variation of MDAs, are not addressed.
2.2	Conversely, for claims with suspected cumulative low doses, does the procedure provide clear guidance in defining worst-case assumptions?	3	
3.0	**Assess the extent to which the procedure accounts for all potential exposures and ensures that resultant doses are complete and based on adequate data in instances where the POC is not evidently clear.**		
3.1	Assess quality of data sought via **interview**:	----	
3.1.1	Is scope of information sufficiently comprehensive?	NA	
3.1.2	Is the interview process sufficiently flexible to permit unforeseen lines of inquiry?	NA	
3.1.3	Does the interview process demonstrate objectivity and is it free of bias?	NA	
3.1.4	Is the interview process sensitive to the claimant?	NA	

3.1.5	Does the interview process protect information as required under the Privacy Act?	NA	
3.2	Assess whether the procedure adequately addresses generic as well as **site-specific data** pertaining to:	----	
3.2.1	Personal dosimeters (e.g., film, TLD, PICs)	NA	
3.2.2	In vivo/In vitro bioassays		
3.2.3	Missing dosimetry data		
3.2.4	Unmonitored periods of exposure		
4.0	**Assess procedure for providing a consistent approach to dose reconstruction regardless of claimants' exposures by time and employment locations.**		
4.1	Does the procedure support a prescriptive approach to dose reconstruction?	3	
4.2	Does the procedure adhere to the hierarchical process as defined in 42 CFR 82.2?	N/A	
5.0	**Evaluate procedure with regard to fairness and giving the benefit of the doubt to the claimant.**		
5.1	Is the procedure claimant favorable in instances of missing data?	2	Does not distinguish work in different areas or job types.
5.2	Is the procedure claimant favorable in instances of unknown parameters affecting dose estimates?	2	See comment 5.1.
5.3	Is the procedure claimant favorable in instances where claimant was not monitored?	3	
6.0	**Evaluate procedure for its ability to adequately account for the uncertainty of dose estimates.**		
6.1	Does the procedure provide adequate guidance for selecting the types of probability distributions (i.e., normal, lognormal)?	5	
6.2	Does the procedure give appropriate guidance in the use of random sampling in developing a final distribution?	5	
7.0	**Assess procedure for striking a balance between the need for technical precision and process efficiency.**		
7.1	Does the procedure require levels of detail that can reasonably be accounted for by the dose reconstructor?	2	Variation of MDAs over time is a necessity, but there is no procedure as to how this information is to be obtained.
7.2	Does the procedure avoid levels of detail that have only limited significance to the final dose estimate and its POC?	5	
7.3	Does the procedure employ scientifically valid protocols for reconstructing doses?	2	Statistical approach is valid, but essential technical details are not addressed.

* Rating System of 1 through 5 corresponds to the following: 1=Never, 2=Infrequently, 3=Sometimes, 4=Frequently, 5=Always. N/A indicates not applicable

5.0 CONCLUSIONS

In conclusion, the statistical methods proposed in ORAUT-RPRT-0044 are based on sound statistical methodologies, and the material is well presented. The proposed methods are an improvement over the regression methods proposed in ORAUT-PROC-0095 when essentially all or most of the data are less-than results, the limit of detection was the same for all samples in the dataset, and **the samples above the limit of detection are randomly spread across workers, job types, and work areas**. Further analyses are needed to evaluate the pattern, time distribution, and worker distribution of the positive results, before a generic coworker daily intake is assigned to the unmonitored worker.

ORAUT-RPRT-0044 does not address the representativeness of the dataset for workers in all work areas and job types. No individual worker analysis was performed, as the report concentrates only on analysis of a collection of analytical results.

The "two-population" method proposed in RPRT-0044 is illustrated by assuming a high CL thought to be reflective of analytical techniques used in the years before the 1980s. Since the data used in this example are 1997 data, it is unlikely that the data reflect working conditions at SRS in the earlier years. The two-distribution approach may only be useful in the analysis of the most recent datasets. RPRT-0044 provides no evidence of the existence of a "two-population" model in datasets from earlier years.

This approach is not claimant favorable for CWs. In general, CWs were not sampled as frequently as NCWs, hence the dataset may not be representative of the distribution of CW exposures. Also, because CWs were sampled less frequently than NCWs, a higher percentage of CWs are expected to require use of the coworker model for dose reconstruction.

REFERENCES

Kumazawa, S. and T. Numakunai, 1981. "A New Theoretical Analysis of Occupational Dose Distributions Indicating the Effect of Dose Limits," *Health Physics*, Vol. 31, No. 3, pp. 465–475. 1981.

Miller, G., T. Little, and R. Guilmette, 2003. "The Application of Bayesian Techniques in the Interpretation of Bioassay Data," *Radiation Protection Dosimetry* Vol. 105, Nos. 1–4, pp. 333–338. 2003.

ORAUT 2005. *Technical Basis Document for the Savannah River Site*, ORAUT-TKBS-0003, Rev. 03, Oak Ridge Associated Universities Team, Cincinnati, Ohio. April 5, 2005.

ORAUT 2006a. *Generating Summary Statistics for Coworker Bioassay Data*, ORAUT-PROC-0095, Oak Ridge Associated Universities Team, Cincinnati, Ohio. June 5, 2006.

ORAUT 2006b. *Y-12 National Security Complex – Occupational Internal Dose*, ORAUT-TKBS-0014-5, Rev. 02, Oak Ridge Associated Universities Team, Cincinnati, Ohio. August 3, 2006.

ORAUT 2007. *Parameters to Consider When Processing Claims for Construction Trade Workers*, ORAUT-OTIB-0052, Rev. 00 PC-1, Oak Ridge Associated Universities Team, Cincinnati, Ohio. January 16, 2007.

ORAUT 2009. *Analysis of Bioassay Data with a Significant Fraction of Less-Than Results*, ORAUT-RPRT-0044, Rev. 00-A, Oak Ridge Associated Universities Team, Cincinnati, Ohio. 2009.

SC&A 2010. *Comparison of ORAUT-OTIB-0052 Plutonium Data for Construction and Non-Construction Workers and the Implications for Coworker Models at Savannah River Site*, SC&A, Inc., Vienna, Virginia, and Saliant, Inc., Jefferson, Maryland. April 2010.

WSRC 1990. *Internal Dosimetry Technical Basis Manual*, WSRC-IM-90-139, Westinghouse Savannah River Company, Aiken, SC. December 20, 1990.