Draft

ADVISORY BOARD ON
RADIATION AND WORKER HEALTH

National Institute for Occupational Safety and Health

DRAFT REVIEW OF PROPOSED ONE PERSON-ONE SAMPLE (OPOS) APPROACH TO COWORKER MODELING

Contract No. 200-2009-28555
SCA-SEC-PR2014-0053, Revision 0

Prepared by

R. Barton
H. Chmelynski
J. Lipsztein
J. Stiver

S. Cohen & Associates
1608 Spring Hill Road, Suite 400
Vienna, VA 22182

February 2014

Disclaimer

S. COHEN & ASSOCIATES: ***Technical Support for the Advisory Board on Radiation & Worker Health Review of NIOSH Dose Reconstruction Program***	Document No. SCA-SEC-PR2014-0053 Effective Date: Draft – February 21, 2014
DRAFT REVIEW OF PROPOSED ONE PERSON-ONE SAMPLE (OPOS) APPROACH TO COWORKER MODELING	Page 2 of 98
Task Manager: _____ Date: _____ John Stiver, MS, CHP	Supersedes: N/A
Project Manager: _____ Date: _____ John Stiver, MS, CHP	Reviewers: A. Makhijani John Stiver John Mauro Steve Marschke

Record of Revisions

Revision Number	Effective Date	Description of Revision
0 (Draft)	02/21/2014	Initial issue

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

Am	americium
AMW	All Monitored Workers
ABRWH	Advisory Board on Radiation and Worker Health
CEDE	committed effective dose equivalent
Cf	californium
Cs	cesium
CTW	Construction Trades Worker
d	day
DOE	U.S. Department of Defense
DTPA	diethylenetriamine pentaacetic acid
EEOICPA	Energy Employees Occupational Illness Compensation Program Act of 2000
FEMP	Fernald Environmental Management Project
GM	Geometric Mean
GSD	Geometric Standard Deviation
h	hour
HHS	Health and Human Services
HIS	Health Information Service
IAEA	International Atomic Energy Agency
ICP-MS	Inductively Coupled Plasma Mass Spectroscopy
IMBA	Integrated Modules for Bioassay Analysis
KPA	Kinetic Phosphorescence Analyzer
L	liter
LCB	Lower Confidence Bound
MCPT	Monte Carlo Permutation Test
MDA	Minimum Detectable Amount
MFP	Mixed Fission Products
MFPG	Mixed Fission Products–Gamma
µg/d	micrograms per day
MPM	Maximum Possible Mean
mrem	millirem

NCRP	National Council on Radiation Protection & Measurements
NCW	Non-Construction Trades Worker
NIOSH	National Institute for Occupational Safety and Health
NIST	National Institute of Standards and Technology
NOCTS	NIOSH OCAS Claims Tracking System
non-CTW	Non-Construction Trades Worker
Np	neptunium
NTS	Nevada Test Site
OPOS	One-Person, One Sample
ORAUT	Oak Ridge Associated Universities Team
OTIB	ORAUT Technical Information Bulletin
POC	probability of causation
Pu	plutonium
ROS	Regression on Order Statistics
RTO	Regression through the Origin
RPRT	Report
SC&A	S. Cohen and Associates
SEC	Special Exposure Cohort
Sr	strontium
SRS	Savannah River Site
U	uranium
UCB	Upper Confidence Bound
WBC	whole-body count
WG	Work Group
WLS	Weighted Least Squares

EXECUTIVE SUMMARY

The concept of aggregating bioassay data to create a one person-one sample (OPOS) statistic for a given time period was extensively discussed during the meeting of the Special Exposure Cohort (SEC) Work Group on September 26, 2013 (ABRWH 2013). SC&A was tasked at that meeting to prepare a report reflecting the various reports that have been prepared by NIOSH and SC&A on the topic, as well as the discussion during the Work Group meeting. SC&A was also asked (NIOSH 2013, p. 3) to be more explicit about its position on the use of the OPOS approach in coworker modeling and comparison of bioassay data of two worker strata.

SC&A agrees with NIOSH that the best approach to coworker modeling is to calculate actual intakes for all monitored workers in the intake regime and to use that distribution for the unmonitored workers. However, NIOSH has stated they do not have the resources to do this. NIOSH also claims that *data dominance* and *correlation* are serious flaws in the current pooled-data coworker modeling approach. As an alternative to the existing pooled-data approach, a *one person-one sample* (OPOS) approach to coworker modeling is recommended by NIOSH as introducing a new scientific credibility to the intake modeling process that does not exist in the pooled-data approach. NIOSH argues that, under certain assumptions, the mean excretion rate is proportional to the intake and serves as a surrogate for the intakes of the individual workers.

The SC&A OPOS review concentrates on the following questions related to the OPOS problem:

(1) Will a large number of incident-related samples from a few workers skew the distributions used for coworker modeling?

(2) When is the mean excretion rate proportional to the intake, and/or does it serve as a surrogate for the intake?

(3) When is it appropriate to use the OPOS approach for coworker modeling and comparing groups of workers?

NIOSH has justified the use of the OPOS approach in comparing two strata of workers to address the problems of data dependence and data dominance. In the former case, a number of bioassay samples following a single intake will be correlated if the radionuclide persists in the body. Hence, the samples are not independent. Data dominance is when a few workers have provided such a large number of samples, as for instance following incidents, that those samples would skew the distributions used for coworker modeling. There is, of course, some overlap between data dependence and data domination, since samples following incidents are not independent.

SC&A has considered the NIOSH responses to its review of *Analysis of Stratified Coworker Datasets*, ORAUT-RPRT-0053 (ORAUT 2012b) carefully. In particular, SC&A reviewed Savannah River Site (SRS) data and *Internal Dosimetry Coworker Data for the Fernald Environmental Management Project*, ORAUT-OTIB-0078, Rev. 2 (ORAUT 2012a) to check whether the issue of data dominance was relevant, since SRS and the Fernald Environmental Management Project (FEMP) are the only sites to which NIOSH has extensively applied the OPOS method.

In a review of bioassays collected at FEMP, "data dominance" was found to be more prevalent than at SRS, particularly when the Code 50 records are included in the analysis. The special records include workers with samples taken daily for a period of time, and are illustrative of how data dominance might influence a coworker model. It is an open question if the coworker model for the unmonitored workers should include these data, and how they should be included. NIOSH has proposed OPOS as one way to deal with this question. A detailed analysis of the FEMP data is included in Appendix C. SC&A found that, in general, NIOSH's assumption of data dominance is not borne out by the data as applied in ORAUT-OTIB-0078, Rev. 2. However, to clearly determine if data dominance is an issue at FEMP, SC&A would need to know which data from HIS-20 were actually used in the OPOS calculations. That information was not available at the time that this report was prepared.

At the same time, there are several examples which show that there are clear cases of data dominance. In such circumstances, the use of the OPOS approach in a limited time period related to a particular incident or special job that resulted in the gathering of an unusually high number of samples may be used to calculate the data point for that worker to be inserted into a distribution of samples otherwise consisting of the pooled data. Each set of samples related to a particular incident would be aggregated in this way and used in the distribution that otherwise would consist of pooled bioassay samples. Such distributions can be used to compare worker strata or to construction of a coworker model, provided other conditions are met.

It should be noted that SC&A's analysis indicates that workers with a large number of samples in a short period of time occur very infrequently. In approximately 95% of the cases where OPOS would be applied at SRS, the workers have no more than five bioassays in the period. The unusual cases may quickly be identified by screening for an unusually large number of bioassays for a worker in any year. Even in these relatively rare cases, there often is no clear evidence of correlation. Thus, the two primary reasons suggested by NIOSH for the use of OPOS have a very limited scope.

The second question above concerns the relationship of the OPOS value to the intake. The decision to recommend the use of OPOS as a general coworker modeling tool, rather than to adjust for occasional incident-driven biases, hinges on whether the mean excretion rate is, in fact, proportional to intake.

The cornerstone of NIOSH's defense of OPOS is the proportional relationship of the mean excretion rate and intake estimate using derived weighted least-squares (WLS) regression. The problem is that the least-squares regression result is only valid to find the intake using the excretion results in urine that are a consequence of that single intake. NCRP Report 164 (NCRP 2013), Appendix A, states the following:

> *This appendix provides a summary of the least-squares method formulas that can be used to derive the intake starting from measurements of activity in bioassay samples.* ***The formulas assume only one intake,*** *no prior knowledge about the magnitude of the intake (i.e., uniform prior in the Bayesian formulation of the intake derivation problem), the biokinetic model and its parameters are known*

> *perfectly, and that all measurements are independent, and properly normalized (e.g., all urine data represent excretion of activity in 24 h).*[emphasis added]

The Integrated Modules of Bioassay Analysis (IMBA) User Manual author notes in "Appendix Section D.8: Using Least Squares Fitting" that the least-squares method is applied to a single intake, and the activity excreted is a result of that intake (James 2005):

Note: The *least squares* fitting method can be used *only* in cases involving a *single intake* – with REAL (explicit) *error values* on each data point, and a *single bioassay quantity*.

When there are multiple intakes, the equations are different and the least-squares procedure does not apply.

NIOSH's justification for the use of a single OPOS value for each worker is based on a method that applies only to excretion results after an intake takes place, and not for excretions resulting from mixed intakes or for urine activities collected from periods of no intakes lumped together with activities from periods with intakes. Despite the NIOSH claim of superior scientific credibility, if OPOS were adopted as a general method for coworker modeling, there is no rationale for applying the weighted least-squares method to infer intakes from urine excretions that are not a consequence of a single intake. It is SC&A's opinion that OPOS is NOT more scientifically credible for general use in coworker modeling, except for the case of fitting bioassay data to a single identified intake in the manner described above.

Regarding question number 3, there are two aspects to the problem: (i) use of OPOS for coworker modeling in general; and (ii) use of OPOS for comparing groups of workers. As noted above, OPOS should be used only in situations when the excretions are known to result from a known intake and in such cases only to aggregate data for those intakes. As a corollary, because the OPOS result is dependent on sampling frequency, OPOS should not be used to dilute bioassay results from periods with known intakes by including results from other periods with no intakes. This caveat is applicable generally to the use of OPOS in coworker modeling.

When OPOS is used to compare groups of workers, additional considerations apply. In this regard, one of our more serious concerns is the use of OPOS to pre-process bioassay data used in hypothesis tests to compare groups of workers when different monitoring protocols were used for the two groups. Although the problem extends beyond this specific application of OPOS, SC&A notes that the protocols by which construction trades workers (CTWs) and non-construction trades workers (NCWs) gave bioassay samples at SRS and other sites may have been different. In those cases, comparisons of the distributions of the data of the two groups of workers are not appropriate. NIOSH has stated that the contractors were not monitored as frequently as the regular workers, and that the contractors' monitoring schedule was incident driven (ABRWH 2013, p. 181). SC&A notes that the problem of comparing bioassay data collected under different monitoring protocols extends to the use of hypothesis tests in general, not only when OPOS is applied to preprocess the data before the tests are applied.

Coworker models for contractors versus regular workers cannot be compared using OPOS, when they have different frequencies of monitoring in the period of no exposure. Examples of illogical results obtained when using OPOS to compare groups of workers with differing monitoring protocols are presented in Section 5.1 of this report. These examples arise due to the dependency of OPOS on sampling frequency, and hence on the monitoring protocol in practice at the time. The OPOS result will depend on the number of monitoring results that the worker had in the year, instead of depending on the number of significant exposures.

Since the OPOS value is a sampling frequency-dependent result, the question of who the coworkers are becomes critical. This is because the protocol used for different workers to collect bioassay samples varies across job types, even leaving aside the CTW versus NCW issue. For instance, at the Nevada Test Site (NTS), the workers most frequently sampled were health physics and related job types. In that case, SC&A found that it was not claimant favorable to assume that the sampled workers were representative of the most exposed workers, even though in most cases, the result would be claimant favorable, but not as a general rule. OPOS aggravates this problem by distorting comparisons and making them dependent on frequency of monitoring.

The OPOS concept should only be applied to those excretion rates related to an identified intake and when many samples were taken from the same worker. In this case, the OPOS value would be calculated only for that period of time when there were a high number of samples. This specific OPOS value (calculated only for the incident-related period, characterized by a large number of samples, all related to the incident) could enter into the lognormal distribution as one result. The excretion rates for the other periods of time should enter the lognormal distribution individually. This approach is in keeping with the NIOSH justification for OPOS: to incorporate into the lognormal distribution results from incidents, when a large number of incident-related samples from a few workers might skew the distribution.

Another major concern for SC&A is that NIOSH intends to use the OPOS statistics as a known value when fitting a coworker model, thus ignoring the uncertainty in this estimate of the mean urine activity for a worker. The introduction of OPOS creates new uncertainties, which may be severe in cases where there are few data points. These uncertainties need to be taken into account whenever OPOS is used. NIOSH has not done so. When the different relative precision of each OPOS value is addressed using weights, the weighted OPOS procedure leads to a coworker model that is very similar to that obtained using the pooled data.

SC&A reaffirms its previous findings regarding the use of OPOS. These include loss of information relating to variability that results from not propagating uncertainties in the averaging process, the requirement that the protocols that were used for collecting bioassay samples from the two strata of workers to be compared be the same, and that, due to loss of variability, the OPOS approach results in a lower value of the upper bound estimate of urinary excretion rates than the former pooled data method. This raises questions regarding claimant favorability in the general application of the OPOS approach if multiple results from the same worker are considered in both the pooled and the OPOS-based coworker models.

In addition to the conceptual problems noted here, there are practical problems with the current procedures for implementing the OPOS approach. In this regard, SC&A has identified several misapplications of OPOS. SC&A notes that NIOSH has not computed OPOS results for SRS or FEMP as per the procedure specified in RPRT-0053 (ORAUT 2012b). NIOSH appeared to acknowledge this problem during the September 26, 2013, Work Group meeting (ABRWH 2013, p. 94 and pp. 265–266). Although NIOSH states that OPOS should be done using the minimum detectable amount (MDA) value for samples that are less than the MDA, this procedure was not always followed.

Finally, the assumption of a year-by-year OPOS analysis without consideration of other factors is a flawed approach to coworker modeling. SC&A recommends that rote use of 1-year analysis periods be modified to match known changes in intake regimes. This may require analysis of periods shorter than 1 year, and the larger size of the pooled dataset will be useful in this regard.

For example if there was a special procedure or an accident in a certain period of time, for example the March–May quarter year, then OPOS would be applicable for the affected workers. Although NIOSH usually applies OPOS on a yearly basis, this period was not specified in ORAUT-RPRT-0053 and other, shorter time periods may be necessary to characterize an identified intake without diluting the result by averaging with periods of non-exposure. The previous coworker models based on the pooled-data approach considered shorter intervals of time. While longer time periods increase available sample size, it obscures the temporal details of worker intakes and exacerbates the problem of sampling frequency dependence.

In conclusion, SC&A recommends that OPOS be used only in a limited context of incident-related samples, with a number of caveats, as explained in the following sections. To use the OPOS concept with justification, it should be applied only for the period of time when there was a large number of urine samples taken after an incident or special job. OPOS should not be applied for the whole year, if this would mix periods of no intake with periods of intake.

Therefore SC&A formulates its principal finding regarding OPOS as follows:

Principal Finding

The use of OPOS on an annual (or other fixed-period) basis as a general matter does not appear to be scientifically justified. *The use of pooled individual bioassay data is recommended despite its known drawbacks. When there is clear evidence of data dominance, the samples related to a particular incident may be averaged to provide a single composite data point to be inserted into the distribution of pooled data, resulting in a "mixed" model.*

SC&A notes that even in this limited context, there is a high degree of uncertainty in the estimated OPOS value due to irregularly spaced collection times, the Regression through the Origin (RTO) hypothesis, the assumption that the variance of the residual error is in direct proportion to the magnitude of intake retention function, and the use of weights inversely proportional to the variance of the measurement.

If accepted, this conclusion would leave us with the pooled-data model with all its known shortcomings as the principal coworker modeling tool. SC&A acknowledges, however, that the substitution of the uncensored results below the MDA in the dataset by the value of the MDA is claimant favorable and should be adopted when deriving intake and doses for unmonitored and monitored workers, regardless of the model that is used to derive intake and doses.

1.0 INTRODUCTION

ORAUT-RPRT-0053 (ORAUT 2012b) introduced the one person-one sample (OPOS) methodology for coworker models. The SC&A review of RPRT-0053 (SC&A 2013) reported three findings which addressed the use of the OPOS methodology:

> ***Finding 3:*** *The OPOS statistic methodology summarizes a worker's exposure by averaging over all urine samples collected during the specified time period. The use of average values does not account for variability of the samples within the time period, and the procedure will result in lower values of the GSD and, hence, the upper bound value used in the coworker model.*
>
> ***Finding 4:*** *The OPOS method must strictly be applied to comparisons where the sampling protocol was the same. Specifically, when there is evidence that the sampling protocol for one group of workers was different than the protocol used for the other group, the tests do not provide a valid comparison. For example, if the monitoring of one group of workers is incident-driven and the other is not, then the OPOS approach is not appropriate for comparing the two distributions.*
>
> ***Finding 7:*** *The statistical tests for comparing two strata require that the samples in each group be independent. If a worker in one group is exposed to radionuclides with long retention in the body and then changes jobs and becomes part of the other group in the same year, the OPOS values are correlated for this worker. This correlation not only violates the assumptions of the tests, but also creates a bias toward a decision of "No Difference" between the two groups.*

The NIOSH response (NIOSH 2013) to SC&A's review of RPRT-0053 contains comments on each of these findings and a rationale for employing the OPOS methodology for coworker models and for comparison of subgroups (strata) of workers. This rationale holds that the primary advantage of OPOS is to "address the problems associated with individuals submitting more than one sample in a year."

This document provides an overview of the issues related to the OPOS findings and responses to the comments in NIOSH 2013. Text quoted directly from NIOSH publications are enclosed in boxes in this report. Additional OPOS-related findings specific to the proposed thorium and neptunium coworker models at SRS are provided in Appendix B. To date, NIOSH has not addressed the SC&A review of these two coworker models. OPOS-related findings for the Fernald uranium coworker model are provided in Appendix C. This appendix represents new material.

2.0 DATA DOMINANCE AND CORRELATION

The NIOSH response to SC&A's review of ORAUT-RPRT-0053 begins the discussion of the one person-one sample (OPOS) issues with the following remarks.

Section 1: Overview
OPOS (p. 3)

In our opinion, the most relevant issue discussed by SC&A is the use of the "one person-one sample" (OPOS) statistic. OPOS was designed to address the problems associated with individuals submitting more than one sample in a year. These problems are:

- Data dominance: a large fraction of the samples being submitted by small fraction of the individuals.
- Correlated data: multiple samples submitted by an individual can be correlated, which greatly complicates the use of statistical tests.

We consider these to be major issues, ones with which the use of the OPOS statistic effectively deals. While more rigorous solutions to these problems may be available, we do not think it is feasible to use them in our situation. SC&A did not comment on the problems of data dominance or correlated data or whether or not the use of OPOS statistics is useful for dealing with them. We would be interested in SC&A's comments on these issues because the review appears to advocate the continued use of individual bioassay results over the use of OPOS statistics. We do not feel that this is a technically viable path forward if one wants to test for differences in strata.

More detailed remarks discussing the motivations for using OPOS are provided further on in the NIOSH response.

Section 3. OPOS DATA (p. 8)

Operational bioassay programs can generate multiple results for an individual in a given period (e.g., a year), which creates a related problem if an individual is involved in an incident and has more (potentially many more[4]) bioassay results than other workers. If these are not accounted for, the problems of correlated data and unequal number of samples per person can violate the assumptions on which the linear regression used to model the data and the statistical tests used to compare strata in the population are based (see Fox 2008, p. 100).

[4] One dataset we observed had 150 bioassay results in a given year, with 100 results from one individual.

The OPOS methodology introduced in RPRT-0053 is applied to the bioassay data used for the coworker model analysis at SRS in ORAUT-OTIB-0081 (ORAUT 2013). Similar arguments are presented in OTIB-0081 for reducing the SRS bioassay dataset by use of the OPOS methodology. The previous NIOSH methodology in ORAUT-OTIB-0075 (ORAUT 2009), which called for use of all individual bioassay samples to fit coworker models, is repudiated in this remark.

OTIB-0081 (p. 16)

In ORAUT-OTIB-0075 (ORAUT 2009), arguments were presented to support the practice of treating a claimant dataset as a simple random sample from the population of all monitored workers. One potential problem posed by using a claimant dataset is that workers involved in incidents usually submit more samples than workers who submit only routine (non-incident-related) samples. This is problematic *because a small number of workers involved in incidents can dominate the claimant sample in a given year through the sheer number of samples submitted and because the samples in the dataset are no longer independent of each other*. At SRS, the small population of workers subject to bioassay testing results in a similar problem. To compensate for the unequal number of samples submitted by workers, the "one person, one sample" (OPOS) technique is used, in which only one result is used for each person for each radionuclide for a given year. The OPOS statistic is calculated using the maximum possible mean methodology (ORAUT 2012a).

Although one of the stated purposes of using OPOS is to reduce the effect of data dominance, it is important to note that NIOSH was also careful to exclude incident-related bioassays for several workers from the OPOS analysis conducted in OTIB-0081. These cases would result in high OPOS values, but the bioassays were removed from the database. One of these cases involved a wound, and it is appropriate that this case be removed, as the coworker models are calculated for inhalation exposure.

OTIB-0081 (p. 19)

Examination of the data revealed occasions during which individuals were involved in incidents that resulted in large intakes and excretions. These data were judged to be unrepresentative of coworkers and were removed. The incidents were:

- One individual was involved in an incident on March 9, 1970. This person's bioassay data were excluded for the remainder of 1970.
- One individual was involved in an incident on March 16, 1972. This person's bioassay data were excluded for the remainder of 1972.
- One individual had a plutonium wound intake on May 8, 1986, which affected the americium bioassay results. This person's bioassay data were excluded for the remainder of 1986.
- Three individuals had false positive results, which were excluded.

If incident-related data are removed prior to analysis, then NIOSH's principal justification for using OPOS (data dominance and correlation) is in question. The incident-related samples are the ones that cause data dominance and OPOS is meant to deal with this problem. SC&A sees no reason for removing incident-related samples AND applying the OPOS procedure.

2.1 EXTENT OF THE PROBLEM

The two arguments presented by NIOSH for data reduction using OPOS are "data dominance" and "correlation." The first argument of data dominance is based on the conjecture that there are "a large fraction of the samples being submitted by small fraction of the individuals." Although this statement suggests there is a serious problem of data dominance, it is a conjecture which has not been shown to have a basis in fact. The fact is that only a *very small percentage of claimants* have multiple bioassays in a short time period in the NIOSH/OCAS Claims Tracking System (NOCTS) SRS database used in the OPOS analysis of many radionuclides, and *almost all bioassay results in the database are from claimants with at most one or two bioassays per quarter*. Data dominance appears to be more significant at Fernald in some periods than was observed for the coworker models proposed for SRS. Under such circumstances, use of the OPOS may be warranted for specific cases.

Similarly, the second argument for data reduction using OPOS is to avoid "correlation." This argument is not based on fact, and again is conjectural. It is not clear whether this statement refers to the possible serial correlation that may occur in samples over time for a given worker, or perhaps to intra-cluster correlation. No evidence is presented that points to the presence of either type of correlation. Statistical methods for establishing the presence of correlation with a high degree of confidence would require many more data points per worker per time period than are available in the claimant database used by NIOSH for stratifying the coworker model.

Lacking firm evidence for the existence of these problems, the NIOSH arguments for using OPOS are largely conjectural. SC&A's position is that conjecture based on anecdotal evidence is not sufficient reason for implementing data reduction using OPOS. If the problems of data dominance and correlation are identified with a high degree of certainty in rare cases, then OPOS data reduction may be appropriate for those cases, but only for the period of time when there is firm evidence that one or both problems exist.

NIOSH has so far applied the OPOS method at SRS and FEMP. SC&A's analysis of the bioassay database used for the SRS OPOS coworker studies did not find evidence of multiple sampling over a short period of time as claimed by NIOSH. The results of that analysis are presented in the following section and Appendix A. As shown in the discussions to follow, there are very few cases with multiple bioassays for an individual in a short period of time.

In an analysis of the NOCTS bioassay database for SRS, the percentage of cases with more than 3 bioassays for a worker in a month was 0.78%, 0.17%, 0.11%, and 1.49% for plutonium, mixed fission products (MFP), and mixed fission products - gamma (MFPG), and americium, respectively. Uranium was somewhat higher, but had less than 5% of cases with more than 3 bioassays for a worker in a month. The frequency of occurrence of a high number of bioassays

for a worker in a quarter is also very low. Appendix C provides a summary of collection frequencies at FEMP.

2.2 NUMBER OF BIOASSAY RESULTS PER CLAIMANT – SRS

The SC&A analysis used claimant bioassay records from the worksheet titled "invitro merged" in the NOCTS urinalysis bioassay data file "NOCTS Data OPOS R20.xlsx." The history and content of this file (Item 11 in the boxes below) is described by NIOSH in the file, "Description of databases used for the SRS internal coworker study_2013-02-15.docx." Note that Item 1 confirms that the claimant bioassay database developed for the SRS coworker study is "complete in that it includes all in-vitro bioassay data for the claim numbers which are included in the spreadsheet." If this is true, then the frequency of claimant bioassay results per worker per time period in the database provides an indication of the extent of the "data dominance" problem in the SRS bioassay data.

> 1. **NOCTS Urinalysis Data** (Spreadsheet is called '*SRS QA_060409 compiled*' (Note-there are 2 worksheets in this spreadsheet)
>
> The SRS site does not have bioassay data prior to 1991 in a format that is easily retrievable, such as an electronic database. For the coworker study, NOCTS bioassay data for the period prior to 1991 was assembled into a single spreadsheet. Much of this data was previously entered into spreadsheets as part of the dose reconstruction process; the remainder was entered specifically to support the coworker study. The format of this spreadsheet is similar to the format commonly used for claim-specific bioassay data files. ***The spreadsheet is complete in that it includes all in-vitro bioassay data for the claim numbers which are included in the spreadsheet. [Emphasis added]***
>
> **…**
>
> 4. **NOCTS Whole Body Count Data** (Spreadsheet is called '*SRS_WBC_WHC_FINAL_Compiled_101811r1 Mike Rev. 1.xlsx'*)
>
> SRS whole body count data is not commonly entered into the claim-specific bioassay data files. Those files typically only note the DOE record page numbers where whole body count (and chest count) data may be found. This [sic] data were needed for the coworker study to evaluate Np intakes between 1970 and 1990 and also to evaluate Cs intakes for all time periods and especially prior to 1991. As with the urinalysis data, the information was compiled into a spreadsheet. The format of the spreadsheet generally matches that used for in-vivo bioassay data at other sites with modifications specific to SRS. Np is not specifically listed as one of the evaluated radionuclides. The information needed to evaluate potential Np intakes based on surrogate data, or data typically used to evaluate other radionuclides, was included. Unlike the in-vitro data, this spreadsheet includes only the information needed for the coworker study and is not a complete record of all the in-vivo data in the DOE record for the claim numbers included in the spreadsheet.
>
> …

9. **Statistical Analysis Instructions** (MSWord file titled "*SRS Coworker Study Statistical Analysis Instruction R37.docx*")

This MSWord document contains the instructions for data extraction and statistical analysis for items 1, 2, 3, 4, 7, and 8. These instructions were also used to generate the spreadsheets from which items 5 and 6 were extracted.
…

11. **OPOS evaluation of NOCTS data** (Spreadsheet titled *NOCTS Data OPOS R20.xlsx*")

This spreadsheet contains the OPOS evaluation of items 1 and 4 above in accordance with the instructions in item 9.

Figures A-11 through A-15 in Appendix A show time plots of the number of bioassay records in the NOCTS SRS database by radionuclide. In general, there are from 50 to 200 bioassay records per quarter for plutonium, MFP, MFPG, and uranium. The smaller number of americium records is the reason additional logbook data are needed for this radionuclide.

The frequency distribution of the number of bioassay results per SRS claimant per month is shown at the left in Table 1. The center columns of Table 1 show the corresponding frequencies for the number of bioassays per claimant per quarter, and annual frequencies are shown on the right. Figures A-1 to A-10 in Appendix A contain plots of the monthly and quarterly frequency distributions shown in Table 1. The data in Table 1 include bioassays for claimants over the entire period from 1950 to 1991.

Table 1. Frequency Distribution of the Number of SRS Bioassay Results per Claimant by Month, Quarter, and Year (1950–1991)

Radio-nuclide	Number of Bioassay Results per Claimant per Time Interval	Length of Time Interval					
		Frequency of Bioassay Results per Month		Frequency of Bioassay Results per Quarter		Frequency of Bioassay Results per Year	
		Frequency	Cumulative %	Frequency	Cumulative %	Frequency	Cumulative %
Pu	1	17,427	91.88%	15,273	86.99%	7,402	62.43%
	2	1,187	98.14%	1,644	96.35%	2,445	83.05%
	3	205	99.22%	346	98.32%	867	90.36%
	4	69	99.58%	148	99.16%	591	95.34%
	5	23	99.70%	47	99.43%	219	97.19%
	6	22	99.82%	41	99.66%	111	98.13%
	7	12	99.88%	17	99.76%	65	98.68%
	8	8	99.93%	14	99.84%	50	99.10%
	9	5	99.95%	7	99.88%	31	99.36%
	10	2	99.96%	9	99.93%	19	99.52%
	>10	7	100%	12	100%	57	100%

Table 1. Frequency Distribution of the Number of SRS Bioassay Results per Claimant by Month, Quarter, and Year (1950–1991)

Radio-nuclide	Number of Bioassay Results per Claimant per Time Interval	Length of Time Interval					
		Frequency of Bioassay Results per Month		Frequency of Bioassay Results per Quarter		Frequency of Bioassay Results per Year	
		Frequency	Cumulative %	Frequency	Cumulative %	Frequency	Cumulative %
MFP	1	6,671	96.35%	6,386	94.47%	3,962	73.17%
	2	223	99.57%	325	99.28%	1,222	95.73%
	3	18	99.83%	33	99.76%	162	98.73%
	4	6	99.91%	10	99.91%	52	99.69%
	>4	6	100%	6	100%	17	100%
MFPG	1	4,388	97.14%	4,264	95.86%	2,528	72.44%
	2	109	99.56%	159	99.44%	806	95.53%
	3	15	99.89%	18	99.84%	114	98.80%
	4	4	99.98%	4	99.93%	30	99.66%
	5	0	99.98%	0	99.93%	6	99.83%
	6	1	100%	2	99.98%	5	99.97%
	>6	0		1	100%	1	100%
Am	1	2,500	90.61%	2,256	87.58%	892	53.25%
	2	185	97.32%	214	95.89%	522	84.42%
	3	33	98.51%	49	97.79%	123	91.76%
	4	16	99.09%	19	98.52%	67	95.76%
	5	3	99.20%	8	98.84%	19	96.90%
	6	6	99.42%	12	99.30%	19	98.03%
	7	1	99.46%	3	99.42%	7	98.45%
	8	3	99.57%	3	99.53%	5	98.75%
	>8	12	100%	12	100%	21	100%
U	1	11,879	82.07%	9,117	76.13%	2,600	41.49%
	2	1,542	92.72%	1,720	90.49%	1,429	64.30%
	3	356	95.18%	450	94.25%	785	76.83%
	4	411	98.02%	207	95.98%	735	88.56%
	5	206	99.45%	63	96.50%	209	91.89%
	6	56	99.83%	72	97.10%	96	93.42%
	7	8	99.89%	49	97.51%	50	94.22%
	8	5	99.92%	41	97.85%	51	95.04%
	9	3	99.94%	28	98.09%	31	95.53%
	10	3	99.97%	32	98.36%	41	96.19%
	11	2	99.98%	38	98.67%	29	96.65%
	12	1	99.99%	46	99.06%	28	97.10%
	13	0	99.99%	59	99.55%	19	97.40%
	14	0	99.99%	34	99.83%	14	97.62%
	15	0	99.99%	9	99.91%	13	97.83%
	16	2	100%	7	99.97%	8	97.96%
	>16	0		4	100%	128	100%

NIOSH proposes to use OPOS on an annual basis. Table 2 summarizes the frequency distribution of bioassay results per year shown at the right in Table 1. The OPOS calculation would not be required for cases with only one bioassay per year in the first row of Table 1.

Table 2 shows the number of OPOS values that would be generated when OPOS is applied on an annual basis to the bioassays shown in the remaining rows of Table 1.

Table 2. Frequency Distribution of the Number of Bioassays in 1-Year OPOS Calculations at SRS

Number of Bioassays in MPM	Pu	MFP	MFPG	Am	U	Total	Cumulative Percentage	Total without Uranium	Cumulative Percentage
1	–	–	–	–	–	–	–	–	–
2	2,445	1,222	806	522	1,429	6,424	56.8%	4,995	65.3%
3	867	162	114	123	785	2,051	74.9%	1,266	81.8%
4	591	52	30	67	735	1,475	87.9%	740	91.5%
5	219	6	6	19	209	459	92.0%	250	94.7%
6	111	4	5	19	96	235	94.0%	139	96.6%
7	65	3	–	7	50	125	95.1%	75	97.5%
8	50	–	–	5	51	106	96.1%	55	98.3%
9	31	–	–	2	31	64	96.6%	33	98.7%
10	19	–	–	3	41	63	97.2%	22	99.0%
11	15	1	1	–	29	46	97.6%	17	99.2%
12	5	1	–	2	28	36	97.9%	8	99.3%
13	7	1	–	1	19	28	98.2%	9	99.4%
14	6	–	–	2	14	22	98.4%	8	99.5%
15	5	–	–	2	13	20	98.5%	7	99.6%
16	3	1	–	2	8	14	98.7%	6	99.7%
17	2	–	–	–	12	14	98.8%	2	99.7%
18	3	–	–	–	9	12	98.9%	3	99.8%
19	1	–	–	–	7	8	99.0%	1	99.8%
20	3	–	–	2	8	13	99.1%	5	99.8%
21	–	–	–	1	3	4	99.1%	1	99.9%
22	2	–	–	–	10	12	99.2%	2	99.9%
23	–	–	–	–	4	4	99.3%	–	99.9%
24	–	–	–	–	8	8	99.3%	–	99.9%
25	1	–	–	1	1	3	99.4%	2	99.9%
>25	4	–	–	3	66	73	100.0%	7	100.0%

Table 2 shows that most OPOS calculations (57%) are performed on workers with only two bioassay results in the year. Seven out of eight (88%) are performed on workers with four or less bioassays per year. Thus, in a typical case, OPOS would be applied to average the results of a small number of quarterly, tri-annual, or biannual bioassays for a worker, rather than using the results individually. When uranium is removed from the totals in Table 2, the results are more compelling. The percentage with only 2 bioassays in the MPM rises to 65%, and the percentage with 4 or less rises to 10 out of 11 cases (91.5%).

Table 3 contains a summary of the annual bioassay results in Table 2. It shows the number of cases with only one sample per year (approximately one-third of cases), and the number of cases where OPOS is required (the other two-thirds). The average number of samples in the maximum possible mean (MPM) when OPOS is applied at SRS ranges from 2.2 samples per MPM for MFP and MFPG, 3.1 for plutonium and americium, and 4.5 for uranium. **On average, there are only 3.4 bioassays included in an MPM calculation at SRS**.

Table 3 also shows the effect on sample size of using OPOS. The total sample size using OPOS is approximately one-half (49%) of the original sample size. The largest reduction in sample size due to use of OPOS (67%) occurs for uranium, which has the largest average number of bioassays in the MPM, and the smallest reduction (25%) occurs for MFP and MFPG, which have the lowest average number of bioassays in the MPM. Reductions for plutonium and americium are near the average at 44% and 50%, respectively.

Table 3. Average Number of Bioassays in 1-Year OPOS Calculations at SRS and Effect on Sample Size

Statistic	Pu	MFP	MFPG	Am	U	Total
(A) Original Sample Sizes	21,293	7,227	4,673	3,325	19,260	55,778
Only 1 Bioassay per Year	7,402	3,962	2,528	892	2,600	17,384
Percent of Bioassays (% of A)	35%	55%	54%	27%	13%	31%
More Than 1 Bioassay per Year	13,891	3,265	2,145	2,433	16,660	38,394
Percent of Bioassays (% of A)	65%	45%	46%	73%	87%	69%
Number of OPOS MPM Values Required	4,455	1,453	962	783	3,666	11,319
Average Number of Bioassays in MPM	3.1	2.2	2.2	3.1	4.5	3.4
(B) New Sample Sizes Using OPOS	11,857	5,415	3,490	1,675	6,266	28,703
Reduction in Sample Size Using OPOS (A minus B)	9,436	1,812	1,183	1,650	12,994	27,075
Percentage Reduction (% of A)	44%	25%	25%	50%	67%	49%

2.2.1 Plutonium

In Table 1, 7 cases were found where a claimant had more than 10 plutonium (Pu) bioassays in the same month. There were 12 cases (0.07% of 17,558 cases) with more than 10 Pu bioassays in the same quarter. These 12 cases were spread over the 40-year time period. In approximately 99.2% of cases, the claimant had 3 or less Pu bioassays in the same month, and 98.3% of cases had 3 or less Pu bioassays per quarter. Nearly 92% of the cases were only 1 sample per month. On a yearly basis, over 95% of cases had 4 or less bioassays per year. In Table 3, on average, only 3.1 plutonium bioassays are used in the annual OPOS MPM calculation at SRS, and use of OPOS results in a 44% reduction in the sample size. It would be difficult to find evidence of correlation in any of these cases, and the effect of data dominance appears minimal, in that it affects only a very small percentage of the data.

The claimant with the second-highest number of samples per quarter is found in the NOCTS SRS plutonium dataset. One claimant working in [redacted] Area had 78 plutonium bioassay results in Q4-[redacted]. These results are a large fraction of the 273 Pu bioassays in that quarter, accounting for 28% of the spike in the total number of plutonium bioassays per quarter seen in [redacted] in the Appendix Figure A-11.

The plutonium bioassay results for this claimant are plotted in Figure 1. The plot indicates that this worker was exposed during an incident in [redacted]. Since we only have NOCTS data, it is not known how many other workers at SRS may have been involved in this incident or if other workers may have had higher plutonium exposures than this claimant. Figure 2 shows the same data on a log scale with a linear regression line. The residuals from this regression should exhibit correlation according to the NIOSH position on correlation. The Durbin-Watson statistic for the log regression is 1.55 and there is evidence of positive serial correlation at the 95% confidence level. It is appropriate to use OPOS in this case with evidence of data dominance and serial correlation. It is not known whether the worker received diethylenetriamine pentaacetic acid (DTPA) treatment following this incident.

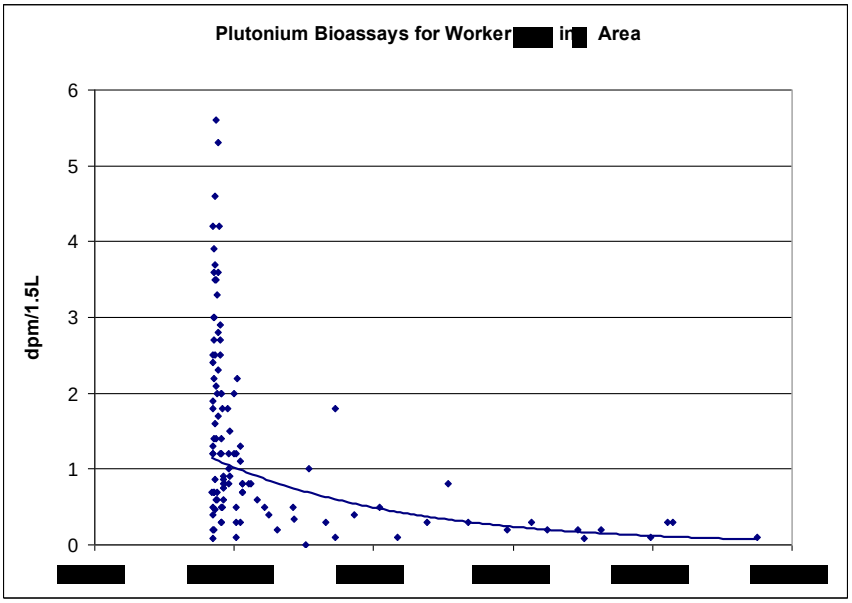

Figure 1. Plutonium Bioassay Results for a Claimant in [Redacted] Area

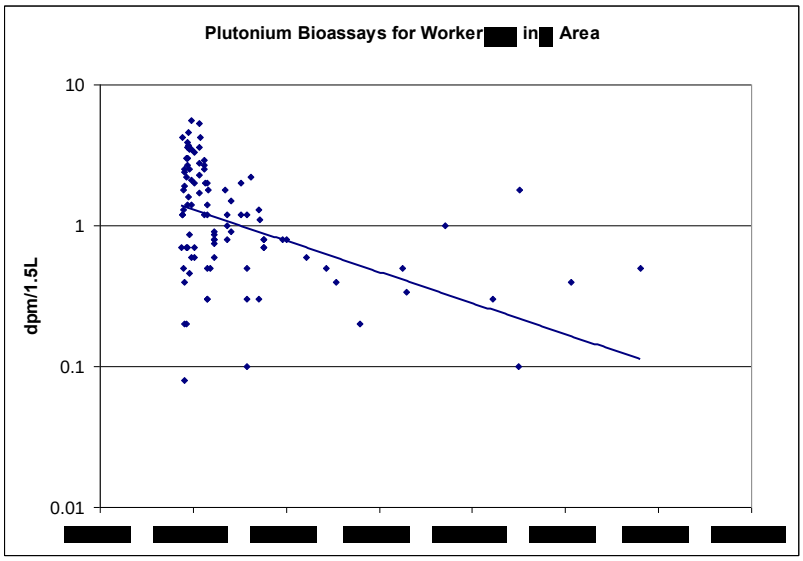

Figure 2. Plutonium Bioassay Results (log scale) for a Claimant in [Redacted] Area with Regression Line

2.2.2 Mixed Fission Products

In the MFP bioassay records shown in Table 1, 12 cases were found where a claimant had 4 or more bioassays in the same month, and 16 cases with 4 or more in the same quarter. In the MFPG bioassays, there were 5 cases with 4 or more bioassays in the same month, and 7 cases with 4 or more in the same quarter. On a yearly basis, 99.7% of cases had 4 or less bioassays per year for MFP and MFPG. The MFP and MFPG records have the lowest proportion of cases with a high number of bioassays per month, quarter, or year. More than 96% of the cases were just 1 sample per month. Table 3 shows an average of only 2.2 MFP or MFPG bioassay results are used in annual OPOS MPM calculation at SRS, and use of OPOS results in a 25% reduction in the sample size.

2.2.3 Americium

For americium, there were 25 cases with 5 or more bioassays in the same month, and 38 cases with 5 or more Am bioassays in the same quarter in Table 1. On a yearly basis, 95.7% of cases had 4 or less bioassays per year for americium. In Table 3, on average, only 3.1 americium bioassays are used in the annual MPM calculation at SRS and use of OPOS results in a 50% reduction in the sample size.

Although these data were not used in the coworker study, it is interesting that the claimant with the highest number of samples per quarter is found in the NOCTS SRS americium dataset. One claimant working in the [redacted] Area had 93 americium bioassay results in Q3-[redacted], with 31 samples per month in June, July, and August of [redacted], averaging 1 per day for 3 months. The americium bioassay results for this claimant are plotted in Figure 3. The plot indicates that this worker was exposed during an incident in the summer of [redacted]. Since

we only have NOCTS data, it is not known how many other workers at SRS may have been involved in this incident or if other workers may have had higher exposures than this claimant.

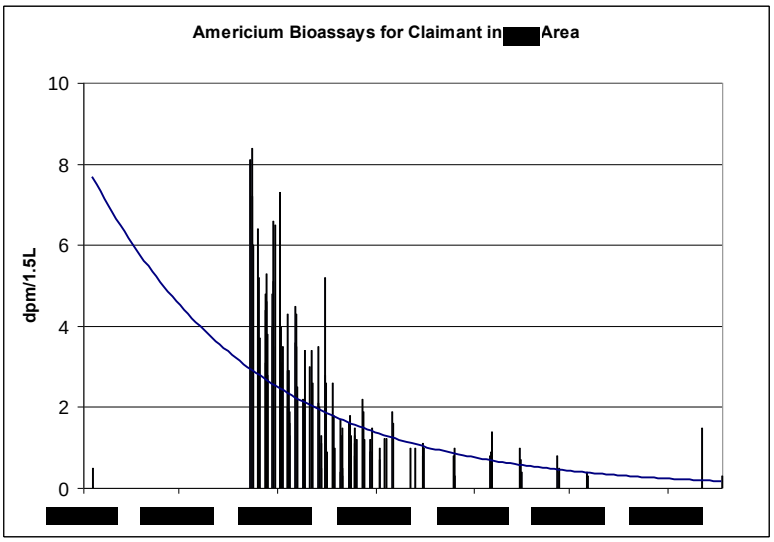

Figure 3. Americium Bioassay Results for a Claimant in [Redacted] Area

The Durbin-Watson statistic for the log regression shown in Figure 4 is 1.09. Again, there is evidence of positive serial correlation at the 95% confidence level. This claimant accounts for approximately one-half of the spike in the total number of americium bioassays per quarter seen in 1970 in Figure A-14. Thus, there is also clear evidence of data dominance in this case. SC&A has no concerns with using OPOS for cases such as this, provided the data are properly compiled. However, this "smoking gun" example of data dominance and correlation was not part of the SRS OPOS analysis, since logbook data were used for americium. It is not known whether this worker received DTPA treatment following the incident.

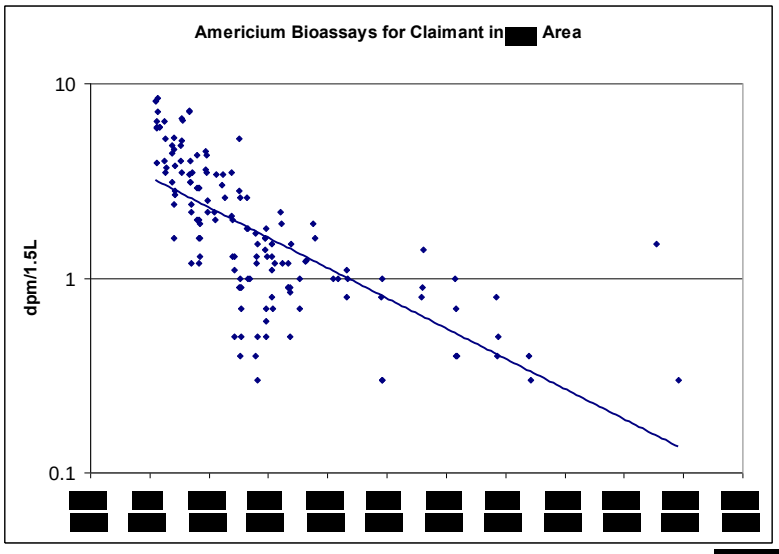

Figure 4. Americium Bioassay Results (log scale) for a Claimant in [Redacted] Area with Regression Line

2.2.4 Uranium

Uranium (U) records had the lowest proportion of cases, with 4 or less bioassays per month or per quarter in Table 1. The percentage of uranium bioassays with 4 or less bioassays in the same month was 98%, versus 99.6%, 99.9%, 99.98%, and 99.1% for plutonium, MFP, MFPG, and americium, respectively. The percentage of uranium bioassays with 4 or less bioassays in the same quarter was approximately 96%, versus 99.1%, 99.9%, 99.9%, and 98.5% for plutonium, MFP, MFPG, and americium, respectively. On a yearly basis, 88.5% of cases had 4 or less bioassays per year for uranium of all types. On average, in Table 3, there are 4.5 bioassays used in the MPM calculations for uranium at SRS, and use of OPOS results in a 67% reduction in the sample size.

3.0 UNCERTAINTY AND VARIABILITY USING OPOS (FINDING 3)

The NIOSH response to the question of lower variability when using OPOS is based on the same two arguments discussed in the previous section; data dominance and independence. NIOSH takes a position that the distribution of OPOS values is a "better" distribution to use for intake modeling; hence, it is "not relevant" that the geometric standard deviation (GSD) of the OPOS distribution is always smaller than that of the individual bioassays. In essence, this point has no more merit than the assumptions on which it is based: that data dominance and correlation are important factors to consider.

Section 2: Responses to Findings
RPRT-0053-3 (p. 6)

The OPOS statistic methodology summarizes a worker's exposure by averaging overall urine samples collected during the specified time period. The use of average values does not account for variability of the samples within the time period and the procedure will result in lower values of the GSD used in the coworker model.

In the presence of **data dominance and dependent data** (see Comment 9), the GM and GSD calculated with individual bioassay measurements do not have familiar statistical properties and are therefore not useful measures of central tendency and variance of the data. The OPOS statistic was adopted in an effort to deal with these major issues. We feel that the use of the OPOS statistic **better** achieves the goal of accurately estimating the intake rates and ultimately the dose to workers than does the use of individual bioassay results. Thus, it **is not relevant** whether or not the OPOS statistics have a higher or lower GSD than the individual data. [*Bold emphasis added.*]

(See Comment 10)

Comment 10 (p. 21)
Referring to page 20 of the SC&A report:

The use of average values does not account for variability of the samples within the time period, and the procedure will result in lower values of the GSD used in the coworker model compared with previous procedures. A GSD must be assigned for the missing dose to a worker in each year, and that GSD should reflect the variability in that worker's exposure during the year. The OPOS GSD measures the variability of average annual dose across workers, and ignores variability for an individual worker within the year.

In the presence of data dominance and dependent data (see Comment 9), the GM and GSD calculated with individual bioassay measurements do not have familiar statistical properties and are therefore not useful measures of central tendency and variance of the data. The OPOS statistic was adopted in an effort to deal with these major issues (on which SC&A did not comment). We feel that the use of the OPOS statistic better achieves the goal of accurately estimating the intake rates and ultimately the dose to workers than does the use of individual bioassay results.

Finding No. 3: The OPOS statistic methodology summarizes a worker's exposure by averaging overall urine samples collected during the specified time period. The use of average values does not account for variability of the samples within the time period and the procedure will result in lower values of the GSD used in the coworker model.

One might infer from this finding that a higher GSD calculated incorrectly is preferable to a lower GSD calculated correctly, perhaps on the basis of "claimant favorability," i.e., higher GSD = higher POC. Below is an excerpt from 42CFR82 that discusses how dose reconstructions should be performed:

"Several commenters requested HHS define what constitutes a 'reasonable estimate' of the radiation doses incurred by an employee. EEOICPA requires the dose reconstruction program to arrive at 'reasonable estimates' of these doses (42 U.S.C. 7384n(d)). HHS interprets this term to mean estimates calculated using a substantial basis of fact and the application of science-based, logical assumptions to supplement or interpret the factual basis. As discussed in the interim final rule, assumptions applied by NIOSH will give the benefit of the doubt to claimants in cases of scientific or factual uncertainty or unknowns."

Thus, if we are presented with multiple, equally valid solutions to a given problem during the process of developing coworker models, we should adopt the solution that gives the benefit of doubt to the claimant. This "claimant favorable" answer is usually taken to be the one that results in the highest dose. The concept of "claimant favorability" is not applicable in the case where there is a solution that is clearly better than the other solutions. More specifically, 42CFR82 does not guide us to adopt an inferior answer simply because it might result in a higher dose than the technically superior answer. Thus, the fact that the GSD will most likely be lower with the OPOS statistics than it is with the individual bioassay results is not relevant because the use of OPOS is a technically superior approach.

Section 3: Detailed Comments

Comment 2 (p. 12)

Referring to page 7 of the SC&A report:

The use of average values does not account for variability of the samples within the time period and the procedure will result in lower values of the GSD used in the coworker model. The OPOS approach represents a significant departure from the previous coworker model methodologies. This change may require re-evaluation of all previous cases with determinations that were based on coworker model estimates.

We agree that the OPOS approach represents a significant change in coworker model methodologies as does the use of stratification – the two are intimately related. Coworker datasets that warrant stratification, and as a result have their coworker models updated to include stratification, will most likely require the re-evaluation of claims based on the unstratified model.

Comment 11 (p. 22)

Referring to page 20 of the SC&A report:

The OPOS methodology does not examine the temporal pattern of individual exposures for longer than one time period.

We are unsure as to exactly what this comment means. In coworker models based on the individual bioassay results, the 50^{th} and 84^{th} percentile excretion rates (or retention) are estimated for each time period – a year for example. With the use of the OPOS statistics the 50^{th} and 84^{th} percentile excretion rates are also estimated for a year. Any subsequent manipulation of the 50^{th} and 84^{th} percentiles is independent of how they were calculated. This is a property of coworker models in general and is the same whether one derives the percentiles with OPOS statistics or the individual bioassay results. So, if the OPOS methodology does not examine the "temporal pattern" of individual exposures for longer than a year then neither does the methodology that uses individual bioassay results.

Comment 13 (p. 24)

Referring to page 21 of the SC&A report:

The answers to these questions are important because the use of OPOS values introduces complications in the subsequent coworker model analyses that rely on these values. OPOS values are not measurements, but are statistics derived from a set of measurements. The OPOS values are averages of a varying number of samples, with a different number for each worker. Since it is an average, each OPOS value has an uncertainty associated with the calculated value.

It is important to realize that all measurements have associated measurement uncertainties and that these uncertainties are not trivial to assess. These facts apply equally to individual bioassay results and to OPOS statistics derived from those individual results. Thus, we find it somewhat inconsistent to take issue with the OPOS statistic for reasons that exist in all measurement data.

Comment 14 (p. 25)

Referring to page 21 of the SC&A report:

A difference in the number of samples available for the workers in each group implies a difference in the uncertainty for the OPOS values for each group. In general, more samples are available for the onsite workers who are part of an ongoing monitoring program. Due to the larger number of samples, the OPOS values for the onsite workers may be measured with greater precision than is available for other groups of workers.

One reason the OPOS statistic was adopted was to give all workers equal weight in the final coworker model – hence the "one person, one sample" moniker. This prevents workers with a larger number of samples per year (onsite workers perhaps) from dominating the coworker model. We find it interesting that SC&A did not offer any comments in this issue or whether or not the use of the OPOS statistic provides any advantages over the use of individual bioassay results.

Since there is uncertainty in the OPOS statistics, and this uncertainty varies from worker to worker and from one group of workers to another, all subsequent analyses based on OPOS values are conducted using heteroscedastic data.

There is measurement uncertainty in all personal dosimetry results (both internal and external – see Comment 13) and all personal dosimetry results are heteroscedastic to some extent. Thus, the issue raised here is not specific to the use of the OPOS statistic.

Comment 16 (p. 27)

Referring to pages 22–26 of the SC&A report:

In Section 3.2, SC&A uses simulations to show that, for a given dataset, a coworker model based on OPOS statistics can have a lower variance (GSD) than a coworker model based on the individual bioassay results and can also give a different estimate of the GM. For example:

The simulation analysis indicates that the OPOS approach results in underestimation of the range of variability across workers reflected in estimates of the GSD and 95th percentile, which are biased low relative to the original samples.

> We agree with these conclusions, which are pretty much what one would have expected to conclude before doing the simulation. However, one might mistakenly infer from this discussion that the model derived from OPOS statistics is somehow "wrong" because it produces estimates of model parameters that are different than the estimates obtained with the individual bioassay results. We feel that we have provided ample technical justification for using the OPOS statistic rather than the individual bioassay results and that any such inference is incorrect.

3.1 SC&A RESPONSE

The issues of *data dominance* and *statistical independence* have merit if the deviations from the standard assumptions of statistical independence and correlation will have an appreciable effect on the distribution of the calculated statistics. However, no quantitative evidence is presented that demonstrates the extent to which the data may contain these problems, other than a few unidentified anecdotal examples. NIOSH has taken a position that a dataset containing any bioassay results that are possibly correlated, or perhaps not proportionally representative of the cohort, should not be used for coworker modeling unless the data are first reduced to a single OPOS statistic for each worker in each analysis period.

Many datasets used in dose reconstruction contain data for only a limited number of workers and/or have incomplete coverage for these workers over time. It is often clear that these records do not contain a complete record for the claimant's or worker's full period of employment, and include many examples where it is known that not all of an individual's test results are included in the database. Given the poor quality of the datasets available for dose reconstruction, at least for some radionuclides at many sites, a strict requirement for ideal statistical properties argues that dose reconstruction using the available retrospective data is not possible. These problems will not be addressed by the use of OPOS.

When only claimant data are available for analysis, only a small fraction of workers at the site are included, and it is not known whether the data are representative. NIOSH claims the data are "complete" in the sense that all available data for the claimants are in the database. This position presumes that NIOSH has access to all data collected at the sites, which cannot be verified. These limitations are not addressed by using OPOS, yet OPOS serves to further reduce the size of the dataset and usefulness of the data. These sources of uncertainty are not reflected in the lognormal distribution estimated from OPOS values.

NIOSH has not addressed the uncertainty in the OPOS estimates derived from a small number of samples from claimants only, with incomplete and/or censored data that are available for any individual in a given year. As shown below in Section 3.3, the magnitude of the uncertainty in the estimates may cover a range as high as a factor of 60. The NIOSH approach ignores the serious problems of uncertainty in the OPOS estimates, but proposes to use the estimates for other claimants' missing values by constructing an OPOS coworker model.

In this regard, SC&A examined in Section 2.1 the data used to compute OPOS statistics for workers at SRS. The results of the analysis show that the data do not support the assumption that data dominance and correlation seriously detract from the analysis of individual bioassay results at SRS. If substantial evidence for data dominance and/or correlation cannot be found, the question remains whether the OPOS statistic is a valid estimate of the intake. The calculation of the OPOS statistic is described in the following section, and the properties of the OPOS estimate are discussed in Sections 3 and 7.

3.2 CALCULATING THE MAXIMUM POSSIBLE MEAN (MPM)

One issue NIOSH fails to address is the uncertainty in the estimated OPOS statistics. To understand the nature of this uncertainty, it is necessary to delve into the specific procedures used to calculate OPOS values from the raw NOCTS bioassay database. The database contains results for a wide variety of radionuclides identified by over 200 symbols, abbreviations and names. SC&A relied on the assignment of individual records to the various radionuclide categories developed by NIOSH: Pu, Sr, Np, and U (all).

The database also contains bioassay records for Am. Although NIOSH now uses logbook urinalysis data for Am for 1972 through 1989, SC&A included the NOCTS Am bioassay results in our review. A large claimant tritium urine bioassay dataset is available from NOCTS for 1954 to 1990, but the OPOS methodology was not used for tritium in the OTIB-0081 SRS coworker study. Therefore, our analysis of the frequency of sampling did not include tritium. A detailed table for classifying radionuclides is shown in Appendix A, Table A-1 of OTIB-0081. The bioassay results in the database are recorded in many different units, and require preprocessing before meaningful comparisons can be made. (Except for incident-related examples presented in Sections 2 and 7, SC&A did not use the bioassay results in this analysis; only record counts.)

The SRS NOCTS bioassay database also contains many records that are not useable for dose reconstruction. The following instructions for the processing of the SRS bioassay data are found in the file listed as Item 9 of the database history in Section 2.2. This file, "SRS Coworker Study Statistical Analysis Instruction R37.docx," contains the following list of instructions for processing the raw bioassay data file in Item 1 above.

Source data and general instructions

- "O:/Coworker Data/Working Files/SRS/Coworker Study/Final reviewed-compiled/SRS QA_060409 compiled.xls" (**NOCTS data**)
  - Do not use records with "N" in column "L," (column titled "use?") except as noted below.
  - Do not use records containing strikethrough font in any cell in columns B through H.
  - Do not use records containing "LIP", L.I.P." or similar in the result, units, or comment fields
  - Do not use records containing "IA" in the result or units fields.
  - Use "one person-one sample" methodology to determine the maximum possible mean of the bioassay data per claimant per analysis period per nuclide.
    - Use the MDA or censoring level for data reported as <MDA.
    - If the analysis period includes all censored data, use the mean result as a censored value for each person.
    - If there are uncensored data during the analysis period, use the mean result as an uncensored result for each person.

Volumes > 1 liter assumed to represent a full day's voiding.
Volumes <= 1 liter should be normalized to 1.5 liter.

The first instruction in the list above (Do not use records with "N" in column "L") eliminates over 10,000 records from the original database of approximately 78,000 bioassay records due to a wide variety of causes. Although some of these omitted records may be "incident-related," the records were not used in the OPOS analysis in OTIB-0081 and are not used here.

The language in the instructions for the MDA calculation should be considered closely. The instructions above include three steps:

- Step 1) Use the MDA or censoring level for data reported as <MDA.
- Step 2) If the analysis period includes all censored data, use the mean result as a censored value for each person.
- Step 3) If there are uncensored data during the analysis period, use the mean result as an uncensored result for each person.

Step 1 starts by replacing all "data reported as <MDA" with the MDA or censoring level. The result of Step 1 is a modified dataset used to calculate the MPM in each analysis period in either Step 2 or Step 3. The instructions consider the case where a data value is reported as a "less-than" value, but do not specifically address the question of what to do when a data value is reported, but it is below the MDA. This case may occur very frequently in some databases and not in others.

The incomplete instructions in Step 1 were a subject of attention at the most recent SEC Work Group meeting. The SC&A position at that meeting was that the MDA should be used for all data with censored values and for all data with entries that are below the MDA. Although numerical values below the MDA (including zero and negative values) may be reported in the data, these entries should not be used to compute the MPM for the period. This part of the procedure was not implemented in several instances examined by SC&A. These shortcomings are noted in Finding 14 for SRS in Appendix B and in Finding C-4 for Fernald in Appendix C. As a result, the MPM for many analysis periods was calculated using both negative values and values below the MDA in the average. The instructions in ORAUT-OTIB-0053 should be revised to avoid misinterpretation.

In Steps 2 and 3 above, the MPM is referred to simply as the "mean result." Step 2 covers the case when all data in the analysis period are below the MDA. In this case, there is a logical basis to "use the mean result as a censored value for each person." If all values for j=1,…, *n* are censored at varied levels L_j, then $x_j < L_j$ and:

$$Average = \sum x_j / n < \sum L_j / n = MPM .$$

If the MPM computed in Step 1 is defined to include the underlined condition above, then the MPM is an average of the MDAs. Hence, if all data are censored in the period, the average of the data (were it known) could be no higher than the average of the MDAs. As its name implies, the MPM is the maximum possible mean in Step 2. Since the MPM is a linear combination of known values (the MDAs), it has no uncertainty. This is appropriate in this case, as the MPM is used for the MDA of a censored OPOS value in Step 2, and an MDA should be a fixed number.

It should be noted that the instructions for Step 2 of the "one person-one-sample" (OPOS) calculations were not implemented as written in the box above. When entries below the MDA are treated as actual values, one would expect only a few cases where all values would be censored and Step 2 would apply. As a result, many OPOS values that should be addressed in Step 2 instead were calculated as uncensored OPOS values in Step 3. During the SEC Work Group meeting on September 26, 2013, NIOSH appeared to agree that the SRS OPOS data are not properly compiled and computed (ABRWH 2013, p. 94 and pp. 265–266).

In Step 3, the MPM from Step 1 is used as the OPOS result for each person in the analysis period. The MPM calculated in Step 1 is either an average of all uncensored values or, more likely, an average of some MDAs and some uncensored values for that worker during the analysis period. In the former case, the MPM is the actual mean of the data. In the latter case, some data values, say x_j for j=1, …, r are uncensored, while others, $x_j = r+1, \ldots, n$, are censored at levels M_j. Then:

$$Average = \frac{\sum_{j=1}^{n} x_j}{n} < \frac{\sum_{j=1}^{r} x_j + \sum_{j=r+1}^{n} M_j}{n} = MPM ,$$

and the "MPM" on the right side of the inequality above is again the maximum possible mean. In this case, the MPM is a linear combination of (*n*-*r*) known values (the MDAs) and *r* random observations, x_j *j*=1, …, r. Unlike Step 2, where the MPM had no uncertainty, in Step 3 calculations, the uncertainty of the MPM must be considered when determining the 84th percentile of the coworker distribution.

Table 4 shows the results of a Crystal Ball simulation showing the range of uncertainty in the OPOS MPM estimate when the number of samples in the MPM ranges from 2 to 12. The first simulation is based on an assumed distribution of bioassay results for the worker that is lognormal with GM=1 and GSD=2.7. The simulated lognormal values were censored to contain 25% non-detects. The MPM of the simulated values is calculated using 20,000 iterations for each sample size. The table shows the point estimate of the MPM, and the lower confidence bound (LCB) and upper confidence bound (UCB) of the 95% confidence interval for the point estimate of the MPM. The confidence intervals are expressed in absolute terms and as a percentage change from the point estimate. A second simulation with GSD=5 is shown at the bottom of the table. The percentage of non-detects is also 25% in this simulation.

Table 4. Point Estimate and 95% Confidence Interval of MPM with 2 to 12 Samples per Analysis Period

GM=1 and GSD=2.72						
Number of Samples in MPM	**Point Estimate of MPM**	**95% Confidence Interval for MPM**				
		Lower Bound		**Upper Bound**		**Ratio=**
		LCB	**%**	**UCB**	**%**	**UCB/LCB**
12	1.70	0.83	-51%	3.4	99%	4.1
10	1.70	0.84	-51%	3.4	99%	4.0
8	1.70	0.78	-54%	3.6	114%	4.7
6	1.70	0.70	-59%	3.9	131%	5.6
4	1.71	0.60	-65%	4.5	163%	7.5
3	1.69	0.54	-68%	4.8	184%	9.0
2	1.70	0.51	-70%	5.7	235%	11.2
GM=1 and GSD=5						
Number of Samples in MPM	**Point Estimate of MPM**	**95% Confidence Interval for MPM**				
		Lower Bound		**Upper Bound**		**Ratio=**
		LCB	**%**	**UCB**	**%**	**UCB/LCB**
12	3.67	0.85	-77%	12.3	235%	14.4
10	3.64	0.85	-77%	12.3	237%	14.4
8	3.73	0.73	-81%	13.7	266%	18.8
6	3.68	0.60	-84%	14.7	301%	24.8
4	3.68	0.45	-88%	16.2	340%	36.0
3	3.67	0.36	-90%	17.9	388%	49.7
2	3.72	0.34	-91%	19.9	435%	58.8

The true mean value of the lognormal distribution is 1.65 with a GSD of 2.72 and 3.65 for a GSD of 5. The point estimates of the MPM are approximately equal to or slightly higher than the true mean value of the lognormal in all cases. The point estimate of the MPM is usually slightly higher than the true mean due to the use of the MDA as a substitute for all non-detects.

The sampling distribution of the MPM calculated using 4 or 12 samples per analysis period is shown in Figure 5. The distributions are approximately lognormal. If the analysis period is one year, then quarterly sampling would yield 4 samples per period and monthly sampling 12 per period. The 95% confidence interval for the MPM point estimate provides a measure of the uncertainty in the MPM estimates. The results of the simulation show the width of the confidence intervals for the MPM.

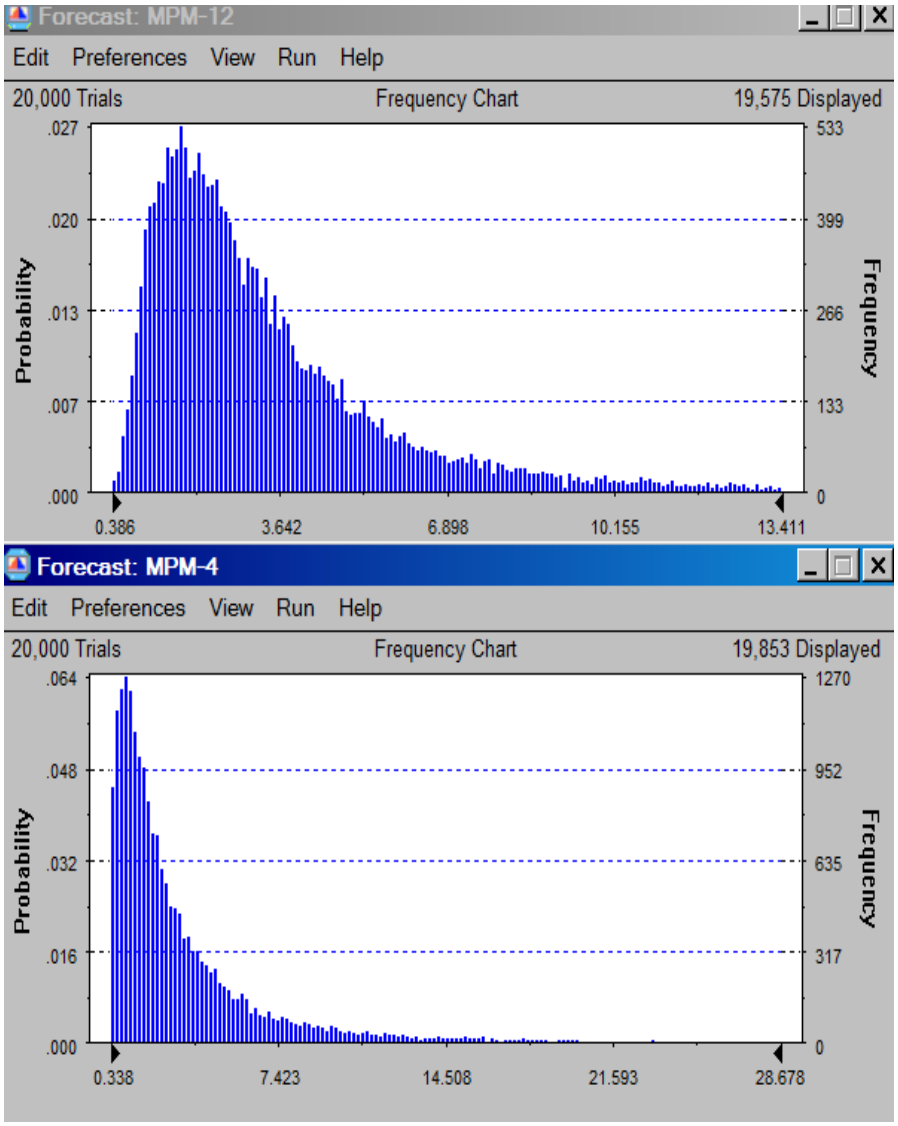

Figure 5. Sampling Distribution of the MPM Average of 4 (lower) or 12 (upper) Lognormal Samples in the Analysis Period

When the GSD is 2.7, the confidence intervals for the MPM range approximately from -50% to +100% if 10 or more samples are included in the MPM. The width of the confidence intervals increases dramatically as the number of samples in the MPM decreases from 10 to 2. With 4 samples in the MPM, the confidence interval width increases to a range from -68% to +184%. With only 2 samples in the MPM, the confidence interval has a range from -70% to +235%. With 4 samples in the MPM, the ratio of the UCB to the LCB is a factor of 4.5/0.6=7.5, and for 2 samples, this ratio increases to a factor of 5.7/0.51=11.2. In summary, with a GSD of 2.72, the uncertainty in the MPM estimate ranges to over a factor of 11.

When the GSD is increased to 5, the confidence intervals are much wider. The confidence intervals range from -77% to +235% when 10 or more samples are included in the MPM. With 4 samples in the MPM, the confidence interval width increases to a range from -88% to +340%. With only 2 samples in the MPM, the confidence interval has a range from -91% to +435%. With 4 samples in the MPM, the ratio of the UCB to the LCB is a factor of 16.2/0.45=36, and for 2 samples, this ratio increases to a factor of 19.9/0.34=58.8. At a GSD of 5, the uncertainty in the MPM estimate spans a range of almost a factor of 60! Uncertainties of these magnitudes cannot be ignored when modeling intakes from the bioassay results. The issue of uncertainty and confidence intervals for the OPOS statistic are discussed in Section 7.7.4.

3.3 MEASUREMENT ERROR

To be precise, the issue here is not "measurement uncertainty," but the error of estimation when measurements are combined into a single OPOS statistic. The MPM is a statistic and hence is only an estimate of the true OPOS value. In Section 3.2, it was shown that the MPM calculated in Step 2 is a fixed value, but the MPM calculated in Step 3 is a linear combination of some fixed values (the MPMs of the censored data) and some random values (the uncensored data). The sampling variance of this statistic should also be calculated as a measure of sufficient accuracy. The sampling variance should also be considered when combining OPOS values into a distribution for the analysis period.

The inclusion of parameter uncertainty is an overriding issue to consider in dose reconstruction. Fitting a distribution to raw measurements is different than fitting a distribution to a collection of statistics. Use of MPMs without consideration of the uncertainty in these estimates will inhibit the propagation of this source of uncertainty when the intake distribution is calculated. Unless parameter uncertainty is included, the GSD used in the probability of causation (POC) calculations will always underestimate the true uncertainty in exposure. This outcome is a direct result of the data reduction that takes place when OPOS is applied. It is not justifiable to introduce this additional significant uncertainty unless there is a large degree of data dominance. In such cases, the uncertainty introduced by the use of the OPOS statistic should be included in the dose calculations. Further discussions of the issue of uncertainty with OPOS are found in Sections 3.2 and 7.4.4.

As noted by NIOSH, the raw data also have measurement uncertainty. This is another source of uncertainty that is not propagated through to the intake calculations. Use of measurement error models should be considered to determine the magnitude of the measurement error.

4.0 OPOS FOR COWORKER MODELING VERSUS OPOS FOR STRATIFICATION

In the following comment, the NIOSH response repeats the concerns over data dominance and correlation, but also draws attention to the difference between using OPOS for coworker models in general versus using OPOS for comparing different groups of workers.

Comment 9 (p. 19)

Referring to Section 3, starting on page 19 of the SC&A report.

One of the main concerns expressed by SC&A about the methodology given in RPRT-0053 appears to be centered around the use of the *one person, one sample* (OPOS) statistic in coworker modeling. For uncensored data, the OPOS statistic for an individual is simply the mean of his bioassay results for a given time period. The OPOS statistic is used in order to deal with two significant issues: dependent coworker data and coworker data dominated by a small number of individuals. Below, we discuss these two problems in more detail and how the OPOS statistic is used to solve these problems.

When multiple bioassays are performed on an individual, the results can be correlated if the individual has had an intake of radioactive material. For example, if an individual has detectable levels of Pu in one urine sample, the next urine sample is also likely to contain Pu. Datasets composed of such *dependent* data usually cannot be analyzed with standard statistical methods, which require *independence* of the data. This issue may have been of marginal importance when all we were interested in was estimating parameters (like intake rates). However, once we start asking if the intake rate of one part of a cohort is different than the intake rate of another part, the issue of data independence becomes critical.

Another problem associated with coworker modeling the bioassay data is that of "data dominance," where a small number of individuals (perhaps even one) submit a significant fraction of the total number of samples collected from the cohort. The resulting coworker model is not representative of the monitored population but instead is dominated by a small number of individuals.

4.1 SC&A RESPONSE

NIOSH notes here that the issues of data dominance and independence are of "marginal importance" when the data are used to estimate a lognormal distribution for the coworker model than they are when distinguishing between groups of workers. Since a comparison of two groups of workers could be conducted simply by comparing the lognormal distributions fitted for each group of workers, it is not clear that there is any difference between these two tasks.

Moreover, the use of the OPOS approach does not necessarily resolve the data dependence problem. For instance, intakes of insoluble uranium will result in dependent data over a period of many years, while soluble uranium may not. But NIOSH makes no attempt to address the issue of prolonged data dependence. Indeed, it is difficult to see how that could be done since the solubility of the intake is generally not known and may, in fact, vary with time due to workers performing different tasks.

The problem here is that NIOSH chooses to use hypothesis tests for comparing the strata. However, because of this choice, the available bioassay data may not meet the requirements for these tests. Then it is necessary to modify the database to fit the choice of methods. SC&A made a number of specific findings in this regard when examining SRS-specific data. They are reproduced in Attachment B of this report.

5.0 SAMPLING PROTOCOL ISSUES (FINDING 4)

NIOSH introduced the concept of OPOS to compare two strata. SC&A's Finding 4 indicates a concern when the OPOS values used to compare two groups of workers are obtained from bioassay results collected under different internal dosimetry monitoring programs, referred to here as "monitoring protocols." This concern applies to the use of hypothesis tests for comparing "apples and oranges." Conclusions based on such comparisons always are suspect. An "apple and oranges" comparison may exacerbate the problem of false positive results (i.e., concluding that the intake potentials of the groups are different when, in fact, they are the same). The problem applies as well to comparisons based on pooled-data approaches.

RPRT-0053-4 (p. 6)

The OPOS method must strictly be applied to comparisons where the sampling protocol was the same. Specifically, when there is evidence that the sampling protocol for one group of workers was different than the protocol used for the other group, the tests do not provide a valid comparison. For example, if the monitoring of one group of workers is incident-driven and the other is not, then the OPOS approach is not appropriate for comparing the two distributions.

We believe that there may be some confusion here concerning the use of the statistical term "sampling protocol." One definition of the term is:[1]

"The sampling protocol is the procedure used to select units from the study population to be measured. The goal of the sampling protocol is to select units that are representative of the study population with respect to the attribute(s) of interest. The sampling protocol deals with how and when the units are selected and how many units are selected."

Thus, the sampling protocol tells how one might select individuals (i.e., a sample) from a population of people with the intent of inferring population parameters from the sample.

In this comment the statistical term "sampling protocol" is incorrectly used as being synonymous with the term "internal dosimetry monitoring program." There is no statistical requirement that all workers be on the same monitoring program in order to use the data to develop a coworker model, as long as the monitoring programs adequately characterize all significant intakes. Further, most sites had *graded* monitoring programs where the frequency and types of bioassay performed were based on the likelihood of the workers having a significant intake of radioactive material.[2] Even today this is standard radiation protection practice, so we would expect the bioassay (i.e., sampling) protocols to be different for different groups of workers.

[1] http://sas.uwaterloo.ca/~rwoldfor/papers/sci-method/paperrev/node40.html

[2] Graded monitoring is also common in external dosimetry programs.

Given all of this, we feel that it is appropriate (for example) to compare intakes calculated from "special" and "task-related" bioassay performed in one group to intakes calculated from "special", "task-related", and "confirmatory bioassay"[3] in another group.

(See Comment 12) [*Note: Comment 12 contains essentially the same response*]

[3] The terms *special*, *task-related*, and *confirmatory* are defined in paragraphs 5.3 and 5.4 in ICRP 1997.

Finding No. 4: The OPOS method must strictly be applied to comparisons where the sampling protocol was the same. Specifically, when there is evidence that the sampling protocol for one group of workers was different than the protocol used for the other group, the tests do not provide a valid comparison. For example, if the monitoring of one group of workers is incident-driven and the other is not, then the OPOS approach is not appropriate for comparing the two distributions.

And the following recommendations offered by SC&A, also on page 22 of their report:

Given the problems introduced by the use of OPOS when there are different sampling protocols for each group, SC&A recommends that:

(1) OPOS values should not be combined into a single lognormal distribution when the sampling protocols for subsets of workers in the group differ
(2) Distributions of OPOS values can be compared only when the sampling protocols are the same for both groups

5.1 SC&A RESPONSE

SC&A would like to be very clear that we are not using the word "sampling" in the above context in the sense of sampling data from a collection of data points. We believe we were quite clear that we are referring to the protocol by which bioassay samples were collected from different groups of workers and hence the protocol by which the data points were obtained. SC&A believes that when two groups are compared through their excretion rates, the OPOS methodology hides important differences in the intake rates of the two groups that are being compared. If one group had urine samples collected because of the occurrence of incidents and the other group collected urine samples from routine monitoring, the OPOS values may show the two groups have different intake potential when, in fact, the intake potentials are the same.

Consider the following hypothetical example:

There are two populations of workers; regular onsite workers and subcontractors. We want to compare bioassay results to determine if the two populations are the same or if there are significant differences. In a given year, there are 10 regular workers in the installation and 10 subcontractors. The regular workers are monitored for uranium every month. The results are very low, all less than MDA (1 activity unit), until suddenly in November there was an accident—a filter rupture—and all workers (regular workers and subcontractors) were exposed. The excretions of all 10 regular workers started to give positive values.

The bioassay results for the regular workers for that year are shown below.

	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec	OPOS
W1	1	1	1	1	1	1	1	1	1	1	70	40	10
W2	1	1	1	1	1	1	1	1	1	1	72	38	10
W3	1	1	1	1	1	1	1	1	1	1	76	46	11
W4	1	1	1	1	1	1	1	1	1	1	74	48	11
W5	1	1	1	1	1	1	1	1	1	1	71	39	10
W6	1	1	1	1	1	1	1	1	1	1	77	45	11
W7	1	1	1	1	1	1	1	1	1	1	70	28	9
W8	1	1	1	1	1	1	1	1	1	1	65	33	9
W9	1	1	1	1	1	1	1	1	1	1	65	21	8
W10	1	1	1	1	1	1	1	1	1	1	66	20	8

The coworker model for the regular workers will be constructed based on the 10 OPOS results: 10, 10, 11, 11, 10, 11, 9, 9, 8, and 8. The arithmetic mean of the OPOS values is 9.7, and the geometric mean is 9.6. The intake rate to be used in the coworker model for the regular workers will be calculated based on an excretion rate of 9.6 activity units in urine. These OPOS values will also be used to compare the two groups of workers.

The subcontractors were not regularly monitored. They were only monitored when there was an indication that they might have been exposed in an incident. In this example, the 10 subcontractors were exposed in the same accident as the regular workers. Their excretion results are very similar to the excretion rates of the workers in November and December:

	Nov	Dec	OPOS
S1	70	40	55
S2	72	38	55
S3	76	46	61
S4	74	48	61
S5	71	39	55
S6	77	45	61
S7	70	28	49
S8	65	33	49
S9	65	21	43
S10	66	20	43

The coworker model for the subcontractors will be constructed based on the 10 OPOS results: 55, 55, 61, 61, 55, 61, 49, 49, 43, and 43. The geometric mean is 53. The intake rate to be used

in the coworker model for the subcontractors will be calculated based on an excretion rate of 53 activity units in urine.

In this example, the intake rates of the coworker model of the subcontractors will be much higher than the intake rates of the coworker model of the regular workers. Yet both groups had the same exposure—the November accident. The only difference was the number of times each worker was monitored before the accident occurred. In reality, the two groups had similar intakes and doses, but the monitoring protocols were different.

This example is related to the fact that the OPOS for the year does not distinguish samples collected following an incident from samples collected in routine monitoring during the year. All results for the year are averaged and the time pattern is ignored. When OPOS is applied to compare two groups of workers, it is necessary that the monitoring protocols be the same for the two groups in order to make a meaningful comparison. The example demonstrates the "apples and oranges" problem noted in the introduction to this section. This concern applies as well to hypothesis test comparisons based on pooled-data approaches. It is expected that the extent of this problem, and the proper resolution thereof, will vary from site to site and from period to period, as monitoring protocols have evolved. Guidelines should be developed for the use of appropriate data when using hypothesis tests for comparing groups of workers.

During the September 26, 2013, Work Group meeting, NIOSH may have agreed that it is not appropriate to mix data of workers who were on an incident-driven monitoring program, though they may have had routine exposure potential with those who were on a routine monitoring program. SC&A notes that the record is not entirely clear on this point (ABRWH 2013):

> *DR. MAKHIJANI: Basically, one of the issues that has concerned us -- and I'm sorry Joyce isn't on the phone, but I will try to represent the situation as best I can for the team -- is that construction workers were thought to be not at routine exposure potential. So, they were only monitored when incidents came to light. But that may not actually be correct. So, it may be a parallel situation or it may not be. We don't have a definitive conclusion about that. But, certainly, we have put this issue on the table in both the reports, the analysis of actual data that we have put on the table for you, more so with the neptunium than with the thorium.*
>
> *DR. NETON: I would agree with you that, if it could be demonstrated the construction workers were on an incident, a certain fraction or a fraction of the construction workers were on an incident-driven bioassay, not part of a regular monitoring program, then that would be not appropriate to incorporate that data into the overall routine monitoring data. I think that is true.*
>
> *DR. MAKHIJANI: But that is what NIOSH has said itself.*
>
> *DR. NETON: Well, I saw that.*
>
> *DR. MAKHIJANI: Right.*

DR. NETON: And it kind of made me take some pause on that comment --

DR. MAKHIJANI: Yes.

DR. NETON: -- because, you know –

DR. TAULBEE: That's not the case, though.

DR. NETON: Okay. If you really have an incident-driven program, there is a separate -- well, okay, I just would agree with Arjun's point that, if there is this sort of dichotomy in monitoring, you know, incident-driven versus routine, I am willing to accept the routine with incident inside of it, sort of a different situation.

DR. TAULBEE: Right. I agree with that.

DR. NETON: That would only tend to bias the results high, but they are still on a routine program. But if you only have incident-driven, then I have got some concern there. [ABRWH 2013, pp. 179–181]

6.0 WORKERS CHANGING JOBS (FINDING 7)

The NIOSH response to Finding 7 argues that the advantages of using OPOS "outweigh" any "minor" problem of non-independence when workers change jobs during the analysis period.

RPRT-0053-7 (p. 9)

The statistical tests for comparing two strata require that the samples in each group be independent. If a worker in one group is exposed to radionuclides with long retention in the body, then changes jobs and becomes part of the other group in the same year, the OPOS values are correlated for this worker. This correlation not only violates the assumptions of the tests, but also creates a bias toward a decision of "No Difference" between the two groups.

First and foremost, we consider the technical benefits realized by using the OPOS statistic to far outweigh relatively rare problems like the one mentioned in this comment. Second, to stratify coworker models one has to be able to assign individuals to specific and meaningful job titles (i.e., develop a *job exposure matrix*). The difficulty in determining an individual's job title, as postulated by SC&A in this comment, is a general problem associated with assembling a job exposure matrix and really has little to do with the use of the OPOS statistic. In fact, the problem raised by SC&A in this comment is an argument for <u>not</u> stratifying a dataset.

We assume that the "violation of assumptions for the tests" mentioned in the comment is the assumption of data independence. The main reason OPOS statistic was adopted was to achieve data independence, which can be grossly violated in the dataset of individual bioassay results.

(See Comment 24)

Comment 24 (p. 40)

Referring to page 37 of the SC&A report:

A second concern with the hypothesis test strategy is that cases may arise when both groups contain the same worker. For example, in the derivative report RPRT-0056 (p. 12), NIOSH states the following [essentially the same passage appears in RPRT-0058 (page 12)]:

Because it was possible for a worker to change jobs during the course of a single evaluated period, it is possible that a worker would have some samples identified as nonCTW and others as CTW in the same period. Therefore, one person might have as many as four different OPOS results, one each for the AMW, CTW, nonCTW, and nonCTW+unk strata.

When the radionuclide is long-lived, the OPOS values generated in each group for that worker will be strongly correlated.

So will the individual bioassay results. This is not a problem that is created by using the OPOS statistic and it can't be solved by using the individual bioassay results. Referring to page 38 of the SC&A report:

Finding No. 7: The statistical tests for comparing two strata require that the samples in each group be independent. If a worker in one group is exposed to radionuclides with long retention in the body, then changes jobs and becomes part of the other group in the same year, the OPOS values are correlated for this worker. This correlation not only violates the assumptions of the tests, but also creates a bias toward a decision of "No Difference" between the two groups.

The correlations in OPOS statistics caused by an individual changing jobs in any given year are considered to be a minor problem because it occurs relatively infrequently. We do not understand the interest SC&A has in this relatively minor contributor to correlations in the data, while at the same time ignoring the relatively significant correlations in individual bioassay results created by individuals submitting multiple samples per year -- a problem the OPOS statistic was meant to address.

Perhaps a more important issue here is that to stratify coworker models one has to be able to assign individuals to specific, unique, and meaningful job titles (i.e., develop a *job exposure matrix*) for all times of employment. The difficulty associated with doing this, as discussed by SC&A in this section is a general problem associated with assembling a job exposure matrix and really has little to do with the use of the OPOS statistic. In fact, the problem raised by SC&A in this section is an argument for not stratifying a dataset and using the standard coworker model.

6.1 SC&A RESPONSE

The NIOSH response to Finding 7 argues that the advantage of using OPOS lies in the conjecture that it solves the "significant" problems of data dominance and correlation, and that this advantage outweighs any "minor" problem of non-independence when workers change jobs during the analysis period. This may be true, but we have no measures of how frequently workers change jobs in the analysis period. Given the very low frequency of sampling shown in Section 2, the two sets of problems may not be that different. SC&A agrees with NIOSH that stratifying coworker models by job titles (and work areas) for all times of employment is the "more important issue here." The difficulties associated with doing analysis by job type were discussed by SC&A as a general problem in coworker modeling. In the SRS NOCTS urinalysis bioassay database, the work area is often known.

It is fundamental that a worker with no data in any time period should be assigned a dose distribution that is based on the bioassays from his/her fellow workers doing the same job, at the same time, in the same work area. In this sense, the coworker model should be thought of as a "k nearest-neighbors" problem. The current approach ignores work area and job title as a means to find a set of similar workers. Using OPOS, bioassays from all workers in the analysis period are combined into a single distribution of averages. No attempt is made to define "coworker" at any finer scale than working in the same time period. Especially at a facility as large and complex as SRS, this one-size-fits-all approach belies the true meaning of the word "coworker." This concern applies as well to hypothesis test comparisons based on pooled-data approaches. It is expected that the extent of this problem will vary from site to site and from period to period.

7.0 INTAKES AND THE OPOS STATISTIC

The NIOSH approach to modeling intakes based on OPOS was described in detail in Comment 9 of their response. The first part of NIOSH Comment 9 was included in Section 3.0 discussion of OPOS and stratification. The second part of Comment 9 addresses the relationship of intakes to the urinalysis bioassay results for an individual worker during the analysis period. The issue here is whether intakes should be calculated based on a single OPOS value for the period. The usefulness of the OPOS approach for modeling intakes was addressed by Dr. Neton at the September 2013 WG meeting:

> *So, what we are saying is an average of an individual worker's bioassay sample is sort of a surrogate for intake. It is directly proportional to their intake. … (p.20)*
>
> *…the average value of a guy's urine data ends up being sort of an indication of picocurie per liter days during that monitoring period of excretion. And, in my opinion, picocurie per liter days of excretion is a very good indicator of intake. It is directly proportional to your intake,* (p.30)

The final part of the NIOSH response Comment 9 expands upon this view that OPOS is a surrogate for the intake. Their response includes a mathematical argument that demonstrates proportionality of the intake with the mean bioassay result under certain assumptions. The SC&A response addresses these arguments and assumptions.

Comment 9 (p. 19) (Continued)

The solution to both of these problems is to model the intakes (or intake rates) rather than the bioassay data. Intakes are independent and if we model the sum of intakes for each individual in a given time period, each person contributes equally to the coworker model. The problem is that it is, in general, not feasible to evaluate each individual's bioassay data in terms of intake.

Given that we cannot use intakes in a coworker model, we chose to use the OPOS statistic as a surrogate for the intake. As shown below, the intake is proportional to the mean of the bioassay data (the OPOS statistic), where the constant of proportionality is the mean of the intake retention fractions.[8] Like an intake, an OPOS statistic is independent and gives each individual in a cohort equal weight in the final model.

While perhaps not the perfect solution to the problems discussed above, we feel that the OPOS statistic is the best available solution, and is undoubtedly better than just modeling the individual bioassay results -- which is the approach that SC&A appears to recommend.

[8] This mean of the intake retention fractions is the part of the intake calculation that is not feasible to determine.

Intakes and the OPOS Statistic

To perform an internal dose coworker model in the most technically correct fashion, we would model the *intakes* of the monitored workers for each year rather than their bioassay data. In IAEA Report 37 [IAEA 2004, pg 22], the equation for the weighted least squares estimate of an intake I is given as:

$$I = \frac{\sum_{i=1}^{n} w_i M_i m_i}{\sum_{i=1}^{n} w_i [m_i]^2}$$

where

n = number of bioassay measurements,
M_i = bioassay measurements,
m_i = intake retention fractions, and
w_i = regression weighting factors.

This is a weighted regression through the origin of the bioassay measurements on the intake retention fractions, and the regression weighting factors are usually taken to be equal to the inverse of the variance of the measurements. If we assume that the variance is proportional to intake retention fraction [Skrable 1994, pg. 442], the weighting factor w_i is given by:

$$w_i = \frac{1}{\left(k \cdot \sqrt{m_i}\right)^2}$$

where k is an unknown constant of proportionality. Under these conditions, the weighted least squares estimate of the intake simplifies to [Skrable 1994, pg. 442]:

$$I = \frac{\frac{1}{n}\sum_{i=1}^{n} M_i}{\frac{1}{n}\sum_{i=1}^{n} m_i} = \frac{\overline{M}}{\overline{m}}$$

which is usually referred to as the "ratio of the means" estimate of intake. Thus, the intake estimate is basically the average of the bioassay results divided by a constant, m, that is determined by the choice of biokinetic models and exposure scenario for that intake. In the case where there are censored bioassay results, one could substitute the OPOS statistic for M to obtain an overestimate of the intake (i.e., the OPOS statistic equals M if all data are uncensored).

7.1 SC&A RESPONSE

In NIOSH's approach to coworker modeling, the time periods are usually on a calendar year or 2-year basis. The OPOS procedure is applied to determine a single average value for each worker with data in the time period. The distribution of OPOS values is then used as the basis for a coworker model for that time period. This response addresses the mathematical and statistical motivations for using OPOS as a summary statistic.

NIOSH has proposed that the intake estimate should be proportional to the mean bioassay for the period. Some authors suggest the use of the weighted least-squares (WLS) method to calculate the intake from multiple bioassay results, all those results being a consequence of this intake. Several of these authors recommend using WLS based on a fundamental assumption that the variance of a bioassay measurement *M* is proportional to the expected value of the magnitude of the bioassay result *E(M)*. Ashley et al. (1992) apply this method for acute intakes using multiple bioassay results related to that intake. Eckerman and Kerr (1999) use the equation:

$$I = \frac{\sum_{j=1}^{N} X_j}{\sum_{j=1}^{N} R_j}$$

to calculate the intakes at Y-12 from bioassay results starting on the date workers started to be exposed to enriched uranium operations:

> *This appendix provides summaries of bioassay data for 15 selected employees who were exposed to enriched uranium following the* ***restart of enriched uranium operations at the Y-12 Plant on June 8, 1998***. ***The bioassay data summarized here starts on June 8, 1998 and ends when the employee was released from restriction.*** *The restrictions were necessary so that follow-up samples in response to results that had exceed Y-12 established trigger levels could be obtained for purposes of dose calculations without interference from additional on-going exposures*. [Emphasis added]

Potter (2005) discusses a single intake and weighting methods for excretion rates related to that intake, including the variance of a bioassay measurement *M* being proportional to the expected value of the magnitude of the bioassay result. However, this paper does not advocate the use of this variance. It recommends, instead, the use of "*average of the slopes:"*

> *The* ***average of the slopes method*** *weights the later data more with respect to the earlier data. Because the models tend to be more representative during longer times post intake,* ***it has been the author's experience that this type of fit is usually more representative.*** [Emphasis added]

Stabin (1997) addresses all activities in urine that result from one single acute intake. Skrable (1988) also discusses a single intake with bioassay data resulting from this intake.

The proportional relationship of the mean excretion rate and intake estimate is a particular case of the WLS regression. As pointed out in NCRP Report 164 (NCRP 2013), the application of

the least-squares regression is only valid to find the intake using the excretion results in urine that are a consequence of that intake. NCRP Report 164 states the following in Appendix A:

> *This appendix provides a summary of the least-squares method formulas that can be used to derive the intake starting from measurements of activity in bioassay samples.* ***The formulas assume only one intake,*** *no prior knowledge about the magnitude of the intake (i.e., uniform prior in the Bayesian formulation of the intake derivation problem), the biokinetic model and its parameters are known perfectly, and that all measurements are independent, and properly normalized (e.g., all urine data represent excretion of activity in 24 h).* [Emphasis added]

NIOSH's justification for the use of a single OPOS value for each worker is based on a method that applies only to excretion results after an intake takes place and not for excretions resulting from mixed intakes, or for urine activities collected from periods of no intake lumped together with activities from periods with intakes.

The NIOSH approach is based on a fundamental assumption that the variance of a bioassay measurement *M* is proportional to the expected value of the magnitude of the bioassay result *E(M)*. This expected value is expressed as the magnitude of the intake *I* times an (unknown) intake retention fraction *R* appropriate for that radionuclide and amount of time elapsed since the intake.

$$Var(M) \propto E(M) = IR$$

Under this and other assumptions, the intake may be shown to be proportional to the unweighted arithmetic average of the bioassay values (i.e., the sample mean) by the application of WLS regression though the origin. To arrive at this result, a series of assumptions is required. The use of least squares is discussed in the following section with assumptions required for WLS through the origin. Applying the formula above in the context of WLS through the origin leads to the desired conclusion.

7.2 WEIGHTED LEAST SQUARES (WLS)

Linear regression for a measured value *x* and one explanatory variable *r* (Figure 6) is described by the equation:

$$x_j = a + Ir_j + e_j$$

Here the symbols x_j represent a set of bioassay measurements after an intake *I*, and the terms r_j are the unknown intake retention fractions[1] associated with each measured value.

[1] When the collection times are known, the notation r_j is a concise notation for the retention rate $R(t_j)$ where t_j is the collection time for measurement x_j. OPOS and the weighted least squares method ignore the ordering of the observations in time. The collection times are an important aspect in the discussion of time-weighted average urine excretion rates in Section 7.

The regression coefficient a is called the intercept (where the regression line crosses the vertical axis), the intake I is the slope of the regression line, and the error terms e_j are both random and homoscedastic. The error term measures the distance of each measurement x_j to the regression line. This distance is called the residual error. If the residual errors are small, the regression is said to have a good fit.

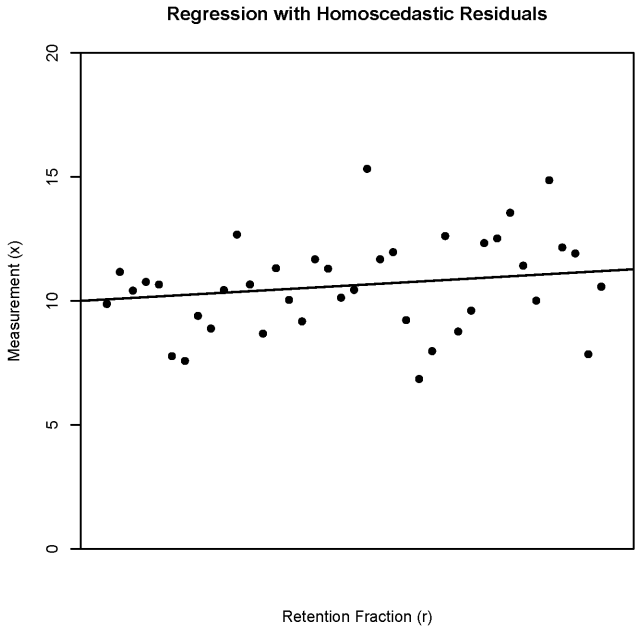

Figure 6. Example of Linear Regression with Homoscedastic Residual Errors

The term "homoscedastic" means that the residual errors have the same variance about the regression line for all values of the intake retention fraction r. In other words, the distance from the measurements to the regression line should be roughly constant over the full range of the independent variable r. If the residuals are not homoscedastic, then residuals are heteroscedastic.

Often there are a priori reasons for believing that the regression line should pass through the origin. Regression through the origin (RTO) with one independent variable is described by a similar equation with no intercept coefficient:

$$x_j = Ir_j + e_j$$

There are many situations when the regression line is expected to pass through the origin, particularly with a non-negative variable. In our case, if the intake retention fraction is 0, then the measured value is expected to be 0.

With RTO smaller values of the intake retention fraction r are typically associated with smaller measured bioassay values x, and larger values of r are associated with larger values of x. With this arrangement of x and r values, the magnitude of the observation is directly proportional to the intake retention fraction r. Heteroscedasticity will occur if the size of the residual error is

related to the magnitude of the observation, and hence to the magnitude of *r*. For example, when using counting statistics the usual statistical model is a Poisson distribution with mean λ. The variance of this Poisson distribution is also λ. Thus small count rates have low variance, and high count rates have high variance. If the RTO model is tenable, small *r* values are associated with small *x* values and hence with small residual errors, while large *r* values are associated with large *x* values and large residual errors.

An example of RTO with homoscedastic residual errors is shown in Figure 7. An example of RTO with heteroscedastic residual errors is shown in Figure 8. This type of heteroscedasticity indicates there is an increasing functional relationship between the standard deviation of the residual and the value of the independent variable. One simple model for the standard deviation of the residual error is a power function:

$$\text{SD} \propto r^p, 0 \leq p \leq 1$$

Figure 9 shows an example of a power function. When the power parameter $p = 0$ standard deviation is a constant for all *r*, so this model includes the homoscedastic RTO model as a special case. This special case is represented by the horizontal line in Figure 9.

If the structure of the data indicates presence of heteroscedasticity, then it is inappropriate to ignore the need to use WLS. WLS regression is applied to correct for the heteroscedasticity. The heteroscedastic RTO (*h*-RTO) model is solved by minimizing the weighted sum of squared residual errors.

The WLS solution is expressed in terms of weighted estimates. The weight for an observation (w_j) is defined as the inverse of the variance of its residual. This definition is designed to restore a condition of equal variance for all residuals. Unfortunately, the RTO residuals will usually have a nonzero mean, because forcing the regression line through the origin is generally inconsistent with the best fit.

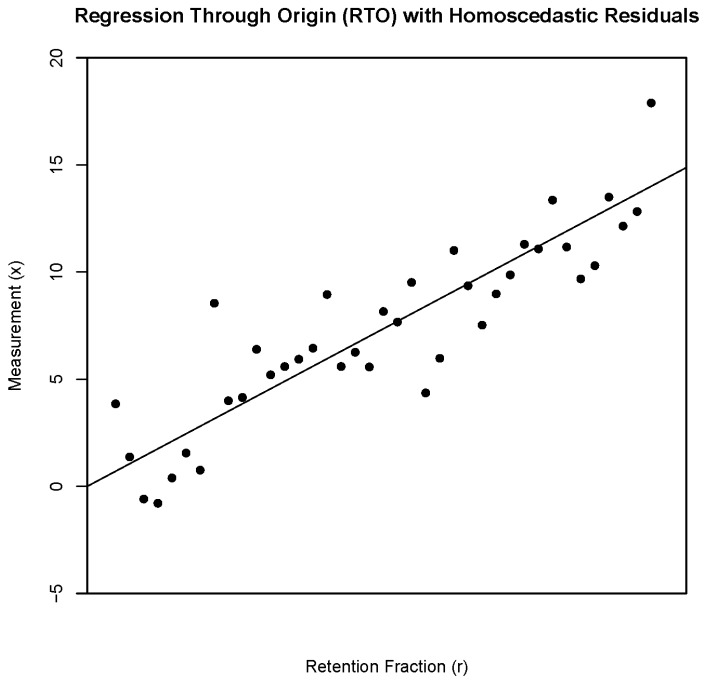

Figure 7. Example of Regression through the Origin with Homoscedastic Residual Errors

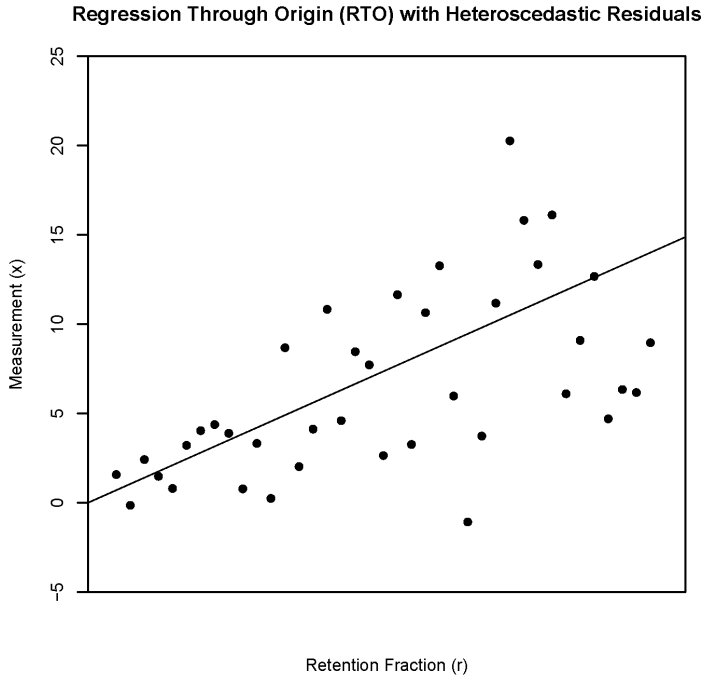

Figure 8. Example of Regression through the Origin with Heteroscedastic Residual Errors

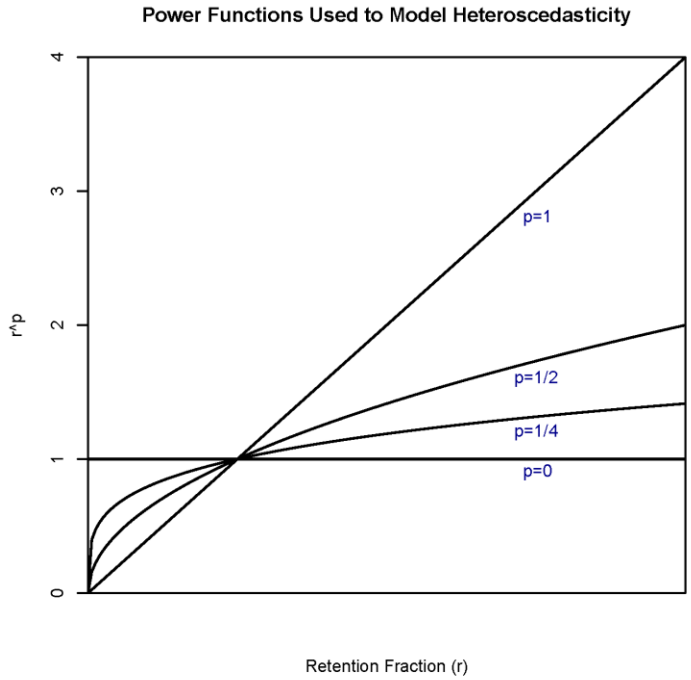

Figure 9. Plots of the Power Function r^p

The proper methods for interpreting regression diagnostics such as R^2 and *F*-statistics for RTO have long been disputed and various software packages will report different results for these diagnostics (Eisenhauer 2003).

With one independent variable, the *h*-RTO model is described by the equation:

$$x_j = Ir_j + r_j{}^p\, e_j$$

where the error term on the right is written as the product of a random and a systematic or nonrandom factor. In this model the factor e_j is random and homoscedastic (i.e., equal in variance, as in the RTO model). The multiplier $r_j{}^p$ in the error term is a systematic factor that is not random. It depends only on the known value of r_j and the power parameter p. The product of these two factors, $r_j{}^p\, e_j$, is random and heteroscedastic.

The variance function describes the way in which the variance of the residual error term differs from observation to observation. If $p > 0$, small observations have small residuals, while large observations have large residuals. If the variance of e_i is σ^2, then the variance of the product error term $r_j{}^p\, e_j$ is $r_j{}^{2p}\, \sigma^2$. In the special case when $p = ½$, the variance of the residual error is in direct proportion to the magnitude of retention factor r_j.

Since the variance is proportional to the square of the power function of *r* shown in Figure 10, the weights are defined to be inversely proportional to the variance of the residual:

$$w_j \propto r_j^{-2p}$$

Although the power function is well-behaved for all values of *r*, the weight function (Figure 10) is very ill-behaved. Unless $p = 0$, the weights asymptotically approach infinity as *r* approaches 0. Thus, very small observations can be assigned very large weight.

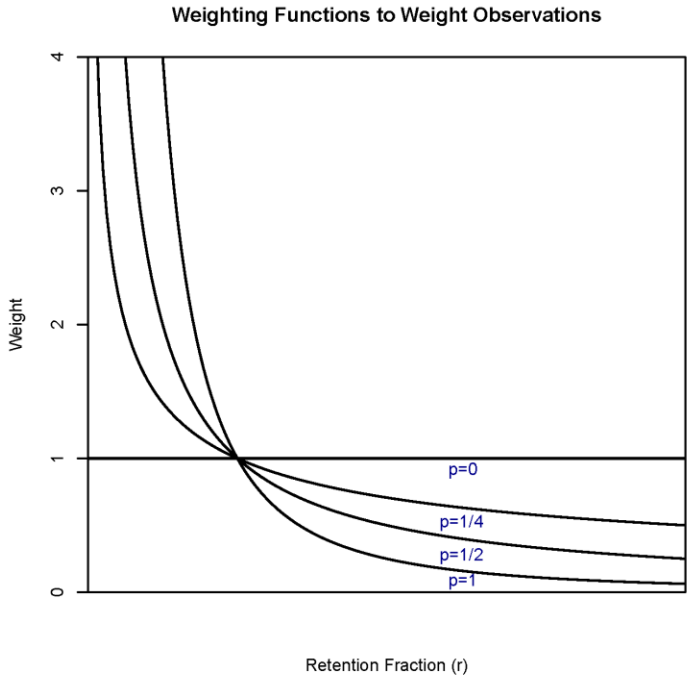

Figure 10. Plots of the Weighting Function r^{-2p}

Despite this undesirable anomaly, the *h*-RTO model with $p = ½$ for the inverse weighting function is selected by NIOSH as the basis to derive a solution of the WLS intake problem introduced in Section 7.1. It should be noted that any other weighting function would not generate a mathematically convenient solution for the OPOS intake calculations proposed by NIOSH. For this same reason, *h*-RTO model with $p = ½$ is used by many authors to solve the multiple bioassay problem. Some of these authors note that alternative weighting schemes (other than $p = ½$) should also be considered, including one with equal weights for all bioassays.

Although heteroscedasticity concerns indicate that the largest observed data points should have less weight in the least-squares regression, it is also necessary to bound the weight assigned to the smallest observations. For example, to avoid giving the smallest observations too much weight, a modified weighting scheme can be used where a constant weight is applied for the first *p* percent of the data points when ranked from smallest to largest. Higher values are assigned weights using a variance function proportion to the value of *r*. With this weighting scheme, the relative weights between data with the smallest *r* values and those with the largest would not be as large.

7.3 USING REGRESSION THOUGH THE ORIGIN TO ESTIMATE INTAKES

As noted in Section 7.1, the variance of a bioassay measurement *M* is assumed proportional to the magnitude of the bioassay result. The expected value of the bioassay result is expressed as the magnitude of the intake *I* times an intake retention fraction *R* appropriate for that radionuclide and amount of time elapsed since the intake:

$$Var(M) \propto IR$$

The final equation used by NIOSH in their response to Comment 9 (found in box in Section 7.0) is only appropriate for a single, determined intake. Thus, the argument for proportionality of the intake with the OPOS value would not be applicable for multiple intakes. This and other limitations of the OPOS approach are discussed further in Section 7.5.

In weighted regression, a standard approach is to use weights assigned inversely proportional to the variance of the measurement:

$$w_j = \frac{1}{kIr_j}$$

Here the constant *k* is selected to make the weights sum to *1*. With this weighting scheme, a high bioassay result is assumed to have a higher variance and hence is assigned a lower weight. Conversely, a low bioassay result is assumed to have a lower variance and hence assigned a higher weight. This weight may be very high, as shown in Figure 10 of Section 7.2. This weighting scheme appears not to be claimant favorable (i.e., high weight on low bioassays and low weight on high bioassays), but does provide a convenient mathematical solution to the problem of estimating the intake from the measurements.

The intake is estimated using WLS regression through the origin. Using the weights given above, the solution takes a simple form. See Appendix D for details.

$$I = \frac{\sum_{j=1}^{N} w_j r_j X_j}{\sum_{j=1}^{N} w_j r_j^2} = \frac{\sum_{j=1}^{N} X_j}{\sum_{j=1}^{N} r_j} = \frac{\frac{1}{N}\sum_{j=1}^{N} X_j}{\frac{1}{N}\sum_{j=1}^{N} r_j}$$

Note the cancellation of the r_j terms when the weights are proportional to $\frac{1}{r_j}$. Hence, in this special case, the intake is found to be proportional to the sample mean of the bioassay results appearing in the numerator of the final expression:

$$\frac{1}{N}\sum_{j=1}^{N} x_j$$

Based on this derivation using WLS through the origin with weights inversely proportional to the variance of the observation, NIOSH has defined OPOS as the sample mean.

NIOSH has made this relationship a cornerstone of their defense of OPOS, but there are several problems with the formula above. The main problem is that least-squares regression result is only valid to find the intake using the excretion results in urine that are a consequence of that intake. NCRP Report 164 (NCRP 2013) states in Appendix A:

> *This appendix provides a summary of the least-squares method formulas that can be used to derive the intake starting from measurements of activity in bioassay samples.* ***The formulas assume only one intake,*** *no prior knowledge about the magnitude of the intake (i.e., uniform prior in the Bayesian formulation of the intake derivation problem), the biokinetic model and its parameters are known perfectly, and that all measurements are independent, and properly normalized (e.g., all urine data represent excretion of activity in 24 h).* [Emphasis added]×

NIOSH is applying the least-squares result indiscriminately to periods of one or more years during which there may be multiple exposure regimes. Applying OPOS on an annual or biennial basis may combine periods of no intake and periods of different intakes. This is not a correct implementation of the WLS result. The IMBA User Manual author notes in *Appendix Section D.8: Using Least Squares Fitting* that the least-squares method is applied to a single intake, and the activity excreted is a result of that intake (James 2005).

Note: The *least squares* fitting method can be used *only* in cases involving a *single intake* – with REAL (explicit) *error values* on each data point, and a *single bioassay quantity*.

When there are multiple intakes, the equations are different and more difficult to solve.

NCRP Report 164, Appendix A, also makes clear that this equation should not be applied if the dataset contains many zero or negative measurement values (NCRP 2013):

> *In practice, data sets often contain many zero or negative measurement values. The lognormal uncertainty model can only be applied for sufficiently positive measurement values, but it is known empirically (Marsh et al., 2007) to be a good representation of actual bioassay uncertainties in this case. To be able to utilize all the measurements, including those near zero or less than zero, the normal/lognormal or additive/multiplicative error model may be used.*

7.4 OPOS AS AN ESTIMATE OF THE MEAN FOR A WORKER

If weighted regression through the origin with the weighting scheme described in the previous sections were deemed acceptable for dose reconstruction, the fundamental question then would be how well the OPOS sample mean performs as an estimate of the true mean for a worker during the period. The answer to this question requires focus on a single worker with a small number of bioassays collected over a period of a year or two. The answer also requires a careful definition of the mean we are trying to estimate.

It is shown in Table 2 of Section 2.1 of the report that in over 90% of the cases where OPOS would be applied at SRS, the typical worker has no more than 5 bioassays in a year. In general, these bioassays are collected at times that are not equally spaced. The time interval between successive bioassay collections from a worker may vary from less than 1 day to a year or more. Figure 11 shows a frequency plot of the number of days between successive bioassays for plutonium for workers at SRS. The data cover a 40-year time frame. The figure shows clear signs of regular quarterly (90-day), biannual (180-day), yearly (365-day), and biennial (730-day) monitoring, but many workers were monitored irregularly.

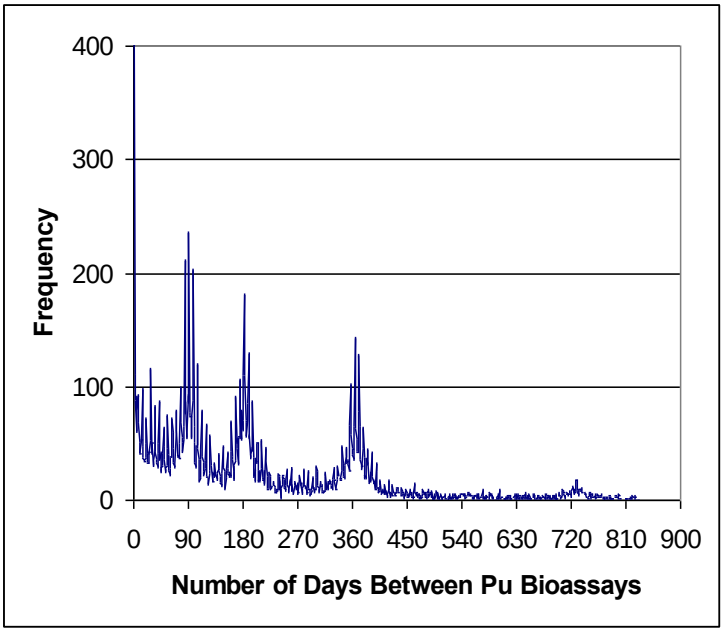

Figure 11. Frequency Distribution of Time Interval in Days between Successive Plutonium Bioassays for Claimants at SRS

7.4.1 Defining the Mean Result

A general approach to the mean value problem posed by OPOS is to construct a time-weighted average using integration. A physical model of the integration process may help motivate this approach. Consider a radionuclide with a half-life much longer than a year.[2] The true mean concentration is defined as the concentration obtained if all of the worker's excretions over the year were pooled together. Included in this mix are days with high activity and others with low activity. We assume the worker is employed for the entire period in question. If the volume excreted is roughly constant at 1.4 L/day, then the mean concentration in the mixture is the time-weighted average of all urine concentrations collected over the year.

To be more specific, let *f(t)* denote a non-negative continuous function which represents the concentration or specific activity of the specified radionuclide in a worker's urine at time *t*.

[2] NIOSH has stated in our technical conference call that they do not give coworker models for short-lived nuclides.

Assume a series of *N* urine bioassay results $f(t_j)$ have been collected at a sequence of times $t_1 \le t_2 \le \ldots \le t_N$. For mathematical convenience, we consider a 1-year analysis period and collection times measured in fractions of a year, where year $t_0 = 0$ is the start of the time period. With this choice of scale, the domain for $f(t)$ has a length of 1, and the area under the curve $f(t)$ is equal to the mean value of the function over the year. If 2-year periods were of interest, then time scale would be measured in fractions of the 2-year period.

Considering the complexity of monitoring intervals shown in Figure 11, there are several ways to define the set of collection times. For some workers the collection times may be on a regular schedule, while other workers may be tested irregularly. The set of collection times $\{t_j\}$ may be represented by either:

(1) A set of N fixed, deterministic collection times or
(2) A set of N randomly selected collection times

The interpretation of OPOS as a mean value differs depending on which assumption is used for the collection times. For example, when the collection times are deterministic and equally spaced, OPOS is equal to the time-weighted average of $f(t)$over the year. If the collection times are deterministic but not equally spaced, then the interpretation of OPOS as a time-weighted average no longer holds. However, if the collection times are truly random (uniformly distributed), then OPOS may be interpreted as a Monte Carlo approximation to the time-weighted average of $f(t)$.

The true mean value of the function $f(t)$is a time-weighted average derived using integration over time. Let $F(T)$ represent the definite integral from time 0 up to time T of the time-varying concentration $f(t)$. The integral may be approximated using, for example, the Riemann sum:

$$F(T) = \int_{t=0}^{T} f(t)dt \cong \sum_{j=1}^{J_T} (t_j - t_{j-1}) \cdot f(t_j).$$

The symbol J_T denotes the index of the last bioassay collection time at or before time T. The sum on the right is a time-weighted sum of all bioassay results for the worker up to time T. The weights are equal to the length of time elapsed since the previous bioassay.

A better approximation to the integral (Crowell and Slesnick 1968) is obtained with the trapezoidal rule:

$$F(T) \cong \sum_{j=1}^{J_T} (t_j - t_{j-1}) \cdot \left(\frac{f(t_j) + f(t_{j-1})}{2} \right).$$

This approach introduces an additional parameter $f(t_0)$ which denotes the concentration at the beginning of the period. In the Riemann approximation, $f(t_0)$ implicitly equals the first bioassay value in the period $f(t_1)$. Another approach would be to use the last bioassay value in the previous time period for $f(t_0)$, if available.

The time-weighted average $\bar{f}$ is equal to the incremental addition made to the integral from the beginning to the end of the calendar year, divided by 1 year:

$$\bar{f} = \frac{\Delta F}{\Delta t} = \frac{F(1) - F(0)}{1\ \text{year}} = F(1).$$

Interpolation may used to estimate the value of the integral F at the end of the calendar year if another bioassay at collection time t_{N+1} is available in the following time period. If not, then the last bioassay result at time t_N would be extrapolated to the end of the year.

7.4.2 Equally Spaced, Deterministic Collection Times

When the bioassays are equally spaced in time and:

$$[(t]_j - t_{j-1}) = 1/N \text{ for } j = 1, \ldots, N ,$$

each bioassay receives the same weight. The weighted Riemann sum is then an unweighted sample mean equal to the OPOS value:

$$\bar{f} \cong \sum_{j=1}^{N} (t_j - t_{j-1}) \cdot f(t_j) = \frac{1}{N} \sum_{j=1}^{N} f(t_j) = OPOS.$$

An example of OPOS as the Riemann sum with equally spaced collection times is shown in Figure 12. In this figure the red curve represents the time-varying urine concentration function *f(t)*. The integral to be evaluated is equal to the area under this curve. Since the integration is over the unit interval from 0 to 1, the integral is equal in value to the mean of *f(t)*. The true mean value is shown in red in the figure. In this example, the curve is a cubic polynomial and the integral may be solved exactly. The OPOS result is in blue. In this case, the OPOS value is equal to the Riemann sum approximation of the time-weighted integral. With a relatively large number of observations and regularly spaced time intervals, both estimates are close to the true value of the time-weighted integral.

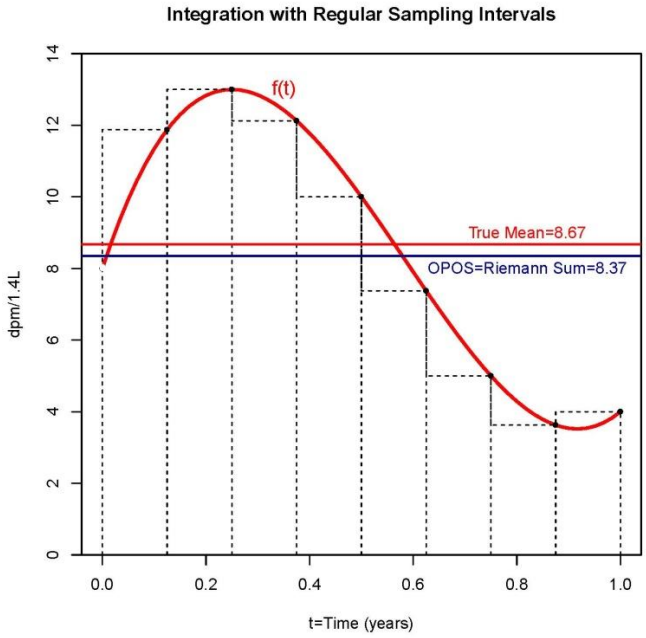

Figure 12. Equivalence of OPOS and Riemann Integral when Collection Times are Regularly Spaced

7.4.3 Unequally Spaced Deterministic Collection Times

When the bioassay collection times are not regularly spaced in time, the weighted Riemann sum applies different weights to each observation. These weights are equal to the length of time elapsed since the previous bioassay. In this case, the unweighted arithmetic average used to compute the OPOS statistic will not equal the Riemann approximation to the integral:

$$\bar{f} = \sum_{j=1}^{N} (t_j - t_{j-1}) \cdot f(t_j) \neq \frac{1}{N} \sum_{j=1}^{N} f(t_j) = OPOS.$$

An example of the difference between OPOS and the Riemann sum with deterministic but unequally spaced collection times is shown in Figure 13. The OPOS result is in blue and the Riemann sum is shown in green. The examples provided in Section 5.1 illustrate some of the problems that arise with irregular monitoring times and situations when the function *f(t)* changes suddenly within the time period.

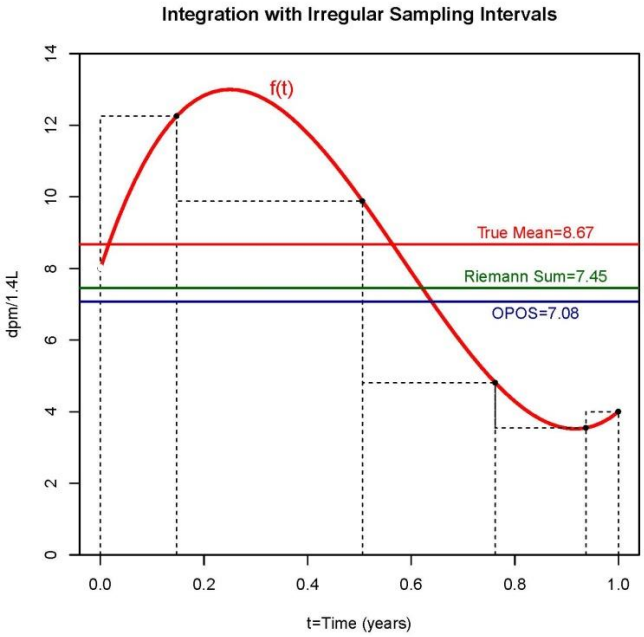

Figure 13. Disparity of OPOS and Riemann Integral when Collection Times are Irregularly Spaced

7.4.4 Random Collection Times

In the previous discussion of deterministic collection times with regular spacing, the OPOS statistic was found to be a reasonable estimate of the mean value of the function. When collection times are assumed to be random and uniformly distributed over the year, all collection times are equally likely.

If the collection times $\{t_j\}$ are random and uniformly distributed over the year, the OPOS statistic is an unbiased estimate of the area under the curve *f(t)*. This result is derived using the basic formula of Monte Carlo integration (Press et al. 1992, Gould et al. 2006):

$$I = \int_{t=0}^{1} f(t)dt \cong \bar{f} \pm \sqrt{\frac{< f^2 > - \bar{f}^2}{N}}$$

Here the sample mean is

$$\bar{f} = \frac{1}{N}\sum_{j=1}^{N} f(t_j) = OPOS$$

and the term $\sqrt{\frac{< f^2 > - \bar{f}^2}{N}}$ is the standard error of the sample mean where

$$< f^2 > = \frac{1}{N}\sum_{j=0}^{N} f^2(t_j)$$

The UCB and LCB of an approximate 95% confidence interval for the integral *I* are derived using the Student *t*-distribution with *N*-1 degrees of freedom:

$$UCB = \bar{f} + t_{N-1}(0.975)\sqrt{\frac{< f^2 > - \bar{f}^2}{N}}$$

and

$$LCB = \bar{f} - t_{N-1}(0.975)\sqrt{\frac{< f^2 > - \bar{f}^2}{N}}$$

The Student *t*-distribution is used, rather than normal distribution, to calculate confidence bounds to account for the fact that the standard error is estimated from the data. It is important to note that the Student *t*-distribution with two or less degrees of freedom has an infinite variance. So if the sample size is *N*=3 or less, percentiles of the *t*-distribution are known for making confidence intervals on OPOS, but the standard error of the OPOS estimate is infinite.

An example of OPOS as Monte Carlo integration is shown in Figure 14. In this figure, the curve *f(t)* is sampled at eight randomly selected times. In this case, the sample mean (OPOS) is the estimate of the area under the curve based on the eight Monte Carlo samples. The upper and lower bounds of the approximate 95% confidence interval for the integral are also shown in this figure. In this case, with eight bioassays in the period, the width of the 95% confidence interval is equal to 61% of the range of the curve.

In almost all cases where OPOS would be applied, the actual number of bioassays for a worker will be smaller than 8, perhaps much smaller. Correspondingly, the confidence intervals for the OPOS sample mean will be wider than those in Figure 14. For example, with 8 bioassays, there are 7 degrees of freedom and the 95% confidence interval is formed using $t_7(0.975) = 2.365$ in the equations above. With 8 bioassays, this *t*-value yields a confidence interval width that is 61% of the range of the data.

Based on the record counts shown in Table 1 in Section 2.2 in over 90% of the cases when OPOS would be applied on an annual basis for plutonium at SRS, the worker has 3 or less bioassays in the year. Consider then a sample size of only 3, yielding 2 degrees of freedom and $t_2(0.975) = 4.303$. The confidence intervals with a sample size of 3 would be 82% wider

$$\left(\frac{4.303}{2.365} = 1.82\right)$$

than with 8 bioassays, spanning more than the full range of the data ($0.61 \times 1.82 = 1.11$). As noted above, the sampling variance of the OPOS estimate with a sample size of 3 is infinite.

Given these extreme measures of uncertainty in the OPOS estimate of the mean, it is inconsistent that NIOSH advocates the use of WLS to address the variance of the observation when relating OPOS to the intake using regression through the origin in Section 7.3, yet does not address the huge uncertainties in the estimated value of the OPOS mean noted here and in the previous

SC&A review of RPRT-0053. The response from NIOSH regarding this (ORAUT 2013) dismissed the concern:

> *It is important to realize that all measurements have associated measurement uncertainties and that these uncertainties are not trivial to assess. These facts apply equally to individual bioassay results and to OPOS statistics derived from those individual results. Thus, we find it somewhat inconsistent to take issue with the OPOS statistic for reasons that exist in all measurement data."*

The uncertainty in the estimated OPOS values indicates a second type of problem with heteroscedasticity encountered when the OPOS data from all workers are fit to a lognormal distribution. As noted in the discussion in Sections 7.2 and 7.3, problems of heteroscedasticity are addressed using WLS. The use of weighted methods for addressing uncertainty in the estimated OPOS values is discussed in Section 7.5.3.

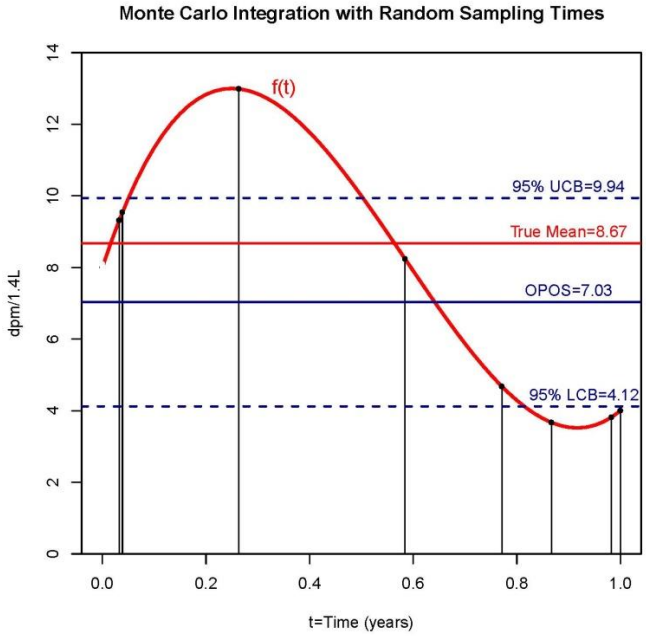

Figure 14. OPOS as Monte Carlo Integral when Eight Collection Times are Randomly Selected, with Upper (UCB) and Lower (LCB) 95% Confidence Bounds for Value of the Integral

7.5 LIMITATIONS OF THE OPOS APPROACH

7.5.1 The Results Apply only to a Single Intake

NIOSH's equation to justify the use of the mean bioassay result only applies to a single intake. The final equation used by NIOSH in their response to Comment 9 (found in the box in Section 7.0) is only appropriate for a single, determined intake. This limitation is related to the use of a single intake term *I* in the weighting equations of Section 7.3 that are used to

demonstrate that OPOS is proportional to the intake. NCRP Report 164 (NCRP 2013) explicitly states in Appendix A:

> *The formulas assume only one intake, no prior knowledge about the magnitude of the intake (i.e., uniform prior in the Bayesian formulation of the intake derivation problem), the biokinetic model and its parameters are known perfectly, and that all measurements are independent, and properly normalized (e.g., all urine data represent excretion of activity in 24 h). The normalization process ensures that measured values are adjusted for all biases and contain all sources of uncertainty related to the measurement process.*

The IMBA User Manual author notes that the least-squares method is applied to a single intake, and the activity excreted is a result of that intake (James 2005, Appendix D.8).

Note: The *least squares* fitting method can be used *only* in cases involving a *single intake* – with REAL (explicit) *error values* on each data point, and a *single bioassay quantity*.

When there are multiple intakes, the equations are different and more difficult to solve.

7.5.2 Alternative Interpretations of the Weighting Term

To justify use of the MPM, NIOSH cites one of the methods for intake estimates from multiple bioassay data, from IAEA Report 37 (IAEA 1997). There is some misunderstanding on the use of the WLS fit, at least what the IAEA Report 37 calls WLS fit.

In the IAEA publication, the WLS fit was one of five methods suggested to interpret bioassay results. IAEA warns that this method should be used with caution. The IAEA Report 37 WLS method uses a weighting factor w_i that will be applied to each bioassay result to allow for the importance of each point, or type of bioassay data. In this case, bioassay data refer to urine bioassay, feces bioassay, in-vivo measurements, etc.

The IAEA points out that the weighting factor (w_i) could be chosen on the basis of subjective assessment of confidence in each point or dataset, or by using some information on the error associated with each measurement. Thus, w_i is not estimated through the equation:

$$w_i = \frac{1}{\left(k \cdot \sqrt{m_i}\right)^2}$$

The weighting term w_i is not, as NIOSH suggests, related to m_i. In general, the weighting term w_i is independent of m_i and is associated with the measurement technique and the laboratory that performed the measurement.

demonstrate that OPOS is proportional to the intake. NCRP Report 164 (NCRP 2013) explicitly states in Appendix A:

> *The formulas assume only one intake, no prior knowledge about the magnitude of the intake (i.e., uniform prior in the Bayesian formulation of the intake derivation problem), the biokinetic model and its parameters are known perfectly, and that all measurements are independent, and properly normalized (e.g., all urine data represent excretion of activity in 24 h). The normalization process ensures that measured values are adjusted for all biases and contain all sources of uncertainty related to the measurement process.*

The IMBA User Manual author notes that the least-squares method is applied to a single intake, and the activity excreted is a result of that intake (James 2005, Appendix D.8).

Note: The *least squares* fitting method can be used *only* in cases involving a *single intake* – with REAL (explicit) *error values* on each data point, and a *single bioassay quantity*.

When there are multiple intakes, the equations are different and more difficult to solve.

7.5.2 Alternative Interpretations of the Weighting Term

To justify use of the MPM, NIOSH cites one of the methods for intake estimates from multiple bioassay data, from IAEA Report 37 (IAEA 1997). There is some misunderstanding on the use of the WLS fit, at least what the IAEA Report 37 calls WLS fit.

In the IAEA publication, the WLS fit was one of five methods suggested to interpret bioassay results. IAEA warns that this method should be used with caution. The IAEA Report 37 WLS method uses a weighting factor w_i that will be applied to each bioassay result to allow for the importance of each point, or type of bioassay data. In this case, bioassay data refer to urine bioassay, feces bioassay, in-vivo measurements, etc.

The IAEA points out that the weighting factor (w_i) could be chosen on the basis of subjective assessment of confidence in each point or dataset, or by using some information on the error associated with each measurement. Thus, w_i is not estimated through the equation:

$$w_i = \frac{1}{\left(k \cdot \sqrt{m_i}\right)^2}$$

The weighting term w_i is not, as NIOSH suggests, related to m_i. In general, the weighting term w_i is independent of m_i and is associated with the measurement technique and the laboratory that performed the measurement.

In addition, the cited IAEA equation is meant to calculate the intake from multiple bioassay measurements related to that particular intake. **The bioassay results from an interval of time in which the worker was not exposed cannot be used together with results from an interval of time in which the worker was exposed, to calculate a mean intake**.

The IAEA Report 37, cited by NIOSH, emphasizes that one of the factors that influence the interpretation of bioassay results is the temporal variation of the intakes of radioactive material. Therefore, SC&A disagrees with NIOSH's assumptions about the weights w_i and hence about its conclusion regarding the relationship about intake to OPOS using the method quoted above.

7.5.3 Uncertainty in the OPOS Statistic

Extrapolating from the simple example in Figure 14, it may be expected that the 95% confidence interval for the true mean value that is estimated by the OPOS statistic will be approximately equal to the full range of the data. With only 2 or 3 samples along the curve *f (t)*, the 95% UCB for the mean may well exceed the maximum observed value.

In this case, there is a trade-off between seeking the "best" estimate of the mean and allowing for uncertainty in that estimate. Using an upper confidence bound on the mean for the OPOS statistic would be more claimant favorable. If the uncertainty range of the OPOS statistic spans the range of the bioassay values for the worker and this uncertainty is properly accounted for in the coworker model, then using the pooled-data coworker modeling approach or the OPOS approach may yield similar GSDs.

As noted in Section 7.2, weighted regression is appropriate when observations have unequal variances. The same arguments apply when fitting a lognormal distribution to the collection of OPOS estimates. In Section 7.3, the WLS regression model was applied to the intake estimation problem for an individual worker using regression through the origin with heteroscedastic error terms. In that application, the intercept term a in the linear regression equation $x_j = a + Ir_j + r_j^p e_j$ was set to 0 to force the regression line through the origin. The ratio of means solution was then obtained by assuming a power coefficient p equal to ½, which implies the weights are inversely proportional to the variance of the error term.

We now consider a second application of the WLS methodology. WLS method is recommended by NIST (NIST 2013) to be used as follows:

> *One of the common assumptions underlying most process modeling methods, including linear and nonlinear least squares regression, is that each data point provides equally precise information about the deterministic part of the total process variation. … This assumption, however, clearly does not hold, even approximately, in every modeling application. …*
>
> *In situations like this, when it may not be reasonable to assume that every observation should be treated equally, weighted least squares can often be used to maximize the efficiency of parameter estimation. This is done by attempting to give each data point its proper amount of influence over the parameter estimates.*

> *A procedure that treats all of the data equally would give less precisely measured points more influence than they should have and would give highly precise points too little influence. …*

In this section, a weighting procedure is applied to estimate a lognormal distribution for using the OPOS estimates for each of the K workers $\{y_1, \ldots, y_K\}$. Each OPOS mean estimate is derived from a different number of bioassays $\{n_1, \ldots, n_K\}$, and the estimate has a sampling variance proportional to the number of values used to compute the mean. The varying number of samples over workers implies that each OPOS estimate is measured with a different precision indicating that the OPOS estimates are heteroscedastic.

Following the NIST recommendations, the weight for each OPOS estimate is assigned in inverse proportion to the sampling variance:

$$w_j \propto n_j .$$

As an introductory example, consider the problem of estimating the mean of a set of OPOS estimates. As shown in Appendix D, the weighted mean is derived as the WLS estimate for the mean using heteroscedastic data. Thus, the weighted mean of the OPOS estimates is the optimal estimate of the population mean. *When the weighted mean of the OPOS estimates is constructed using weights proportional to the number of observations, the result is equal to the mean of the pooled data*:

$$\hat{\mu} = \frac{\sum_{k=1}^{K} w_k y_k}{\sum_{k=1}^{K} w_k} = \frac{\sum_{k=1}^{K} n_k y_k}{\sum_{k=1}^{K} n_k} = \frac{\sum_{k=1}^{K} n_k \sum_{j=1}^{n_k} \frac{x_{j,k}}{n_k}}{\sum_{k=1}^{K} n_k} = \frac{\sum_{k=1}^{K} \sum_{j=1}^{n_k} x_{j,k}}{N}$$

Where $x_{j,k}$ is the j^{th} bioassay result for worker k and the total number of samples in the pooled data is:

$$\sum_{k=1}^{K} n_k = N$$

This example demonstrates that the proper use of weights for the OPOS estimates leads to an estimate for the population mean equal to the result obtained from the pooled data, but requires use of a more complicated algorithm.

Each OPOS estimate also should be weighted by the number of bioassays used to compute that OPOS estimate when fitting a lognormal distribution for the coworker models. The weights should be used when Regression on Order Statistics (ROS) are applied to estimate a lognormal distribution for coworker modeling, and for comparisons of groups of workers which involve the use of lognormal distributions as in the Monte Carlo Permutation Test. The weights are to be assigned in inverse proportion to the variance of the OPOS estimate, and thus in direct proportion to the number of bioassays used in the OPOS calculation for that worker.

ROS using weighted OPOS estimates is very similar to ROS when using unweighted OPOS values. The only difference is that a weighted ranking is constructed by summing the weights

assigned to the OPOS values when calculating plotting points. In weighted ROS, the rank of an observation is replaced by the sum of the normalized weights.

An example of the weighted ROS is shown in Figure 15. The example was constructed using 3 years of plutonium bioassay results at SRS. The years selected include multiple instances of workers with a relatively large number of bioassays in the period. The OPOS estimate was computed for each worker using the MPM procedure, and weights were assigned to each estimate in proportion to the number of bioassays used in the calculation.

Figure 15 shows a comparison of ROS plots for:

(1) Pooled data (ROS using all bioassays)
(2) Unweighted OPOS (ROS using unweighted OPOS values)
(3) Weighted OPOS (ROS using weighted OPOS values)

The weighted OPOS distribution tends to follow that of the pooled data, except in the extreme upper percentiles. In the highest percentiles, the weighted OPOS and unweighted OPOS distributions have the same end points. In most other regions of the plot, the unweighted OPOS distribution is lower than the other two distributions by approximately an order of magnitude in this example.

The 95th percentile of the distributions is marked by a vertical line near the center of the figure. The pooled data and the weighted OPOS distributions have approximately the same 95th percentile, while the 95th percentile of the unweighted OPOS distribution has a 95th percentile that is lower by approximately an order of magnitude.

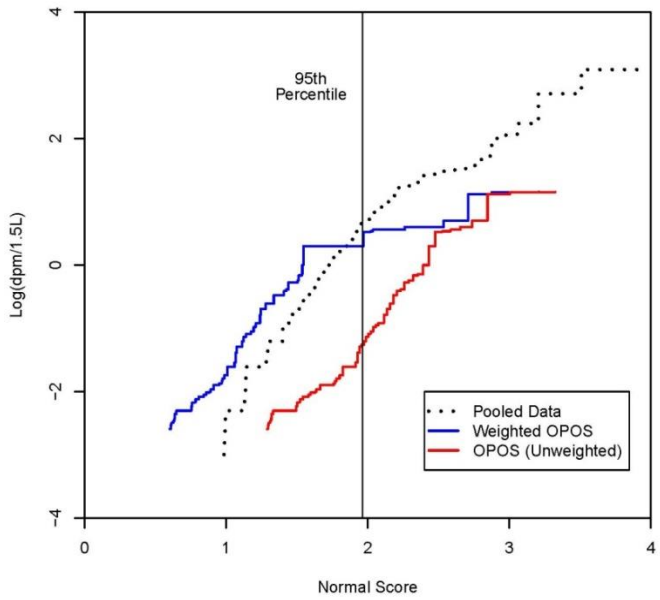

Figure 15. Pooled Data, Weighted OPOS and Unweighted OPOS Distributions

This example indicates that the OPOS procedure, with the proper use of weights to reflect the precision of each OPOS estimate, leads to a coworker model that is very similar to what one

would obtain using the pooled data, except that the pooled data model has many fewer assumptions and calculations. The use of OPOS without proper weighting of the estimates leads to a coworker model that is applied with questionable logic (no weights) and defended with a "Rube Goldberg" string of questionable assumptions (WLS and RTO). As indicated in the quote from NIST (2013): "A procedure that treats all of the data equally would give less precisely measured points more influence than they should have and would give highly precise points too little influence." In the case of the coworker model, the workers with the higher number of bioassays tend to be the workers with the higher bioassay values, and these are the workers who are given *too little influence* in the coworker model. As a result, the OPOS procedure leads to a coworker model that is not claimant favorable.

Given these concerns, the pooled-data model is recommended as a much more transparent approach to coworker modeling, despite its known drawbacks. If the OPOS procedure is adopted, SC&A recommends the use of weights that reflect the precision of each OPOS estimate.

7.5.4 Non-parametric Approaches to Coworker Model Comparisons

NIOSH has argued that there should be only one statistic for each worker. There is no requirement that OPOS must be the sample mean. In fact, OPOS is not the sample mean when non-detects are present in the data, but instead a "Maximum Possible Mean."[3]

In the Monte Carlo Permutation Test (MCPT), workers are selected in each random draw, and then are assigned their OPOS values for estimating the lognormal distribution. SC&A prefers an approach where the value assigned to the worker in that draw is a randomly selected bioassay value collected from that worker during the analysis period. Random selection of a value will permit the full range of values for each worker to have a chance of selection, while assigning equal weight to each worker.

[3] Use of this name for the OPOS statistic should cease, as it is quite deceptive. The true maximum possible mean is reflected by the Upper Confidence Bounds for the sample mean discussed in Section 7.4.4.

REFERENCES

42 CFR 82, 2002. *Methods for Radiation Dose Reconstruction under the Energy Employees Occupational Illness Compensation Program Act of 2000*; Final Rule, Federal Register/Vol. 67, No. 85/Thursday, May 2, 2002, p. 22314.

ABRWH 2013. *Transcript of the meeting of the Special Exposure Cohort Work Group*, Advisory Board on Radiation and Worker Health, September 26, 2013. On the web at http://www.cdc.gov/niosh/ocas/pdfs/abrwh/2013/wgtr092613.pdf.

Ashley, J.C., J.M. Barber, L.M. Snapp, and J.E. Turner, 1992. Martin Marietta Energy Systems, Inc., for the U.S. Department of Energy, *Technical Basis for the Internal Dosimetry Program At the Y-12 Plant*" Y/DQ-40, March 2, 1992.

Crowell, R. H. and W. E. Slesnick, 1968. *Calculus with Analytic Geometry*, Norton & Company Inc., New York, New York.

Eckerman, K.F. and G.D. Kerr, 1999. *Y-12 Uranium Exposure Study, Appendix E ACTLITE Code listing*, Oak Ridge National Laboratory, Lockheed Martin Energy Systems, Inc., ORNL/TM-1999/114, August 5, 1999.

Eisenhauer, J. G., 2003. "Regression through the Origin," *Teaching Statistics*, Volume 25, No. 3.

Fox, J., 2008. *Applied Regression Analysis and Generalized Linear Models*, Sage Publications, London, England.

Gould, H., J. Tobochnik, and W. Christian, *2006. An Introduction to Computer Simulation Methods: Applications to Physical Systems, 3rd edition*, Addison-Wesley, Boston, Massachusetts.

IAEA 2004. Methods for Assessing Occupational Radiation Doses Due to Intakes of Radionuclides. *Number 37 in Safety Report Series*. International Atomic Energy Agency, 2004.

ICRP 1997. Individual Monitoring for Internal Exposure of Workers, ICRP Publication 78 *Annals of the ICRP* 27 (3-4).

James, A.C., 2005. *IMBA Expert™ OCAS/ORAU-Edition User Manual, Appendix D: Example Bioassay Cases*, ACJ & Associates, Inc., Richland, Washington.

Marsh, J.W., E. Blanchardon, C.M. Castellani, A.D. Desai, M.D. Dorrian, C. Hurtgen, V. Koukouliou, M.A. Lopez, A. Luciani, M. Puncher, A. Andrasi, M.R. Bailey, V. Berkovski, A. Birchall, Y. Bonchug, H. Doerfel, I. Malatova, A. Molokanov, and H. Ratia. 2007. "Evaluation of scattering factor values for internal dose assessment following the IDEAS guidelines: Preliminary results," Radiat. Prot. Dosim. 127(1-4), 339-342.

NCRP 2013. *Uncertainties in Internal Radiation Dose Assessment*, Report No. 164, National Council on Radiation Protection & Measurements, Bethesda, Maryland, 2013.

NIOSH 2013. *Response to SC&A Comments on ORAUT-RPRT-0053,* Rev. 0, August 2013.

NIST 2013. NIST/SEMATECH e-Handbook of Statistical Methods, accessed 11/01/2013 http://www.itl.nist.gov/div898/handbook/pmd/section1/pmd143.htm

ORAUT 2009. *Use of Claimant Datasets for Coworker Modeling, ORAUT-OTIB-0075,* Rev. 00, Oak Ridge Associated Universities Team, Cincinnati, Ohio. May 2009.

ORAUT 2012a. *Internal Dosimetry Coworker Data for the Fernald Environmental Management Project*. ORAUT-OTIB-0078, Rev. 2. Oak Ridge Associated Universities Team, Cincinnati, Ohio. December 12, 2012.

ORAUT 2012b. *Analysis of Stratified Coworker Datasets*, Rev. 1. ORAUT-RPRT-0053, Oak Ridge Associated Universities Team, Cincinnati, Ohio. July 2012.

Potter, C.A., 2005. "Internal dosimetry: A review." *Health Phys*. 88(6): 565–578.

Press, W.H., S.A. Teukolsky, W.T. Vettering, and B.P. Flannery, 1992. *Numerical Recipes in FORTRAN – The Art of Scientific Computing 2nd Ed.*, Cambridge University Press, New York, New York.

SC&A 2013. *Draft Review of ORAUT-RPRT-0053: Analysis of Stratified Coworker Datasets,* Rev. 1, April 2013.

Skrable, K.W., G.E. Chabot, C.S. French, and T.R. LaBone, 1988. "Intake Retention Functions and Their Applications to Bioassay and the Estimation of Internal Radiation Doses," *Health Physics Journal*, Volume 55, No. 6, 1988.

Skrable 1994. Skrable, K.W. et al., *Estimation of Intakes from Repetitive Bioassay Measurements*, Chapter 19 in Internal Radiation Dosimetry, Otto G. Raabe, editor. (Madison: Medical Physics Publisher), 1994.

Stabin, Michael G., 1997. "Demystifying Internal Dose Calculations," www.doseinfo-radar.com/demystify.doc, substantial portions of the text for this document were taken from the author's publication "Internal Radiation Dosimetry," in *Nuclear Medicine,* Henkin et al., eds, Mosby-Year Book, St. Louis, Missouri. 1996.

.

APPENDIX A: ANALYSIS OF NOCTS SRS BIOASSAY RECORD DATABASE

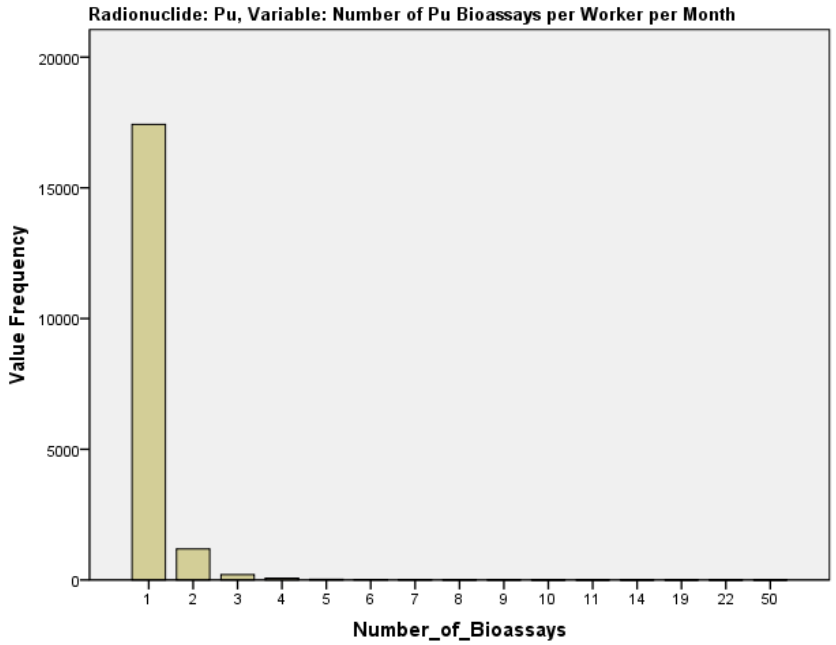

Figure A-1. Frequency Distribution of the Number of Plutonium Bioassay Results per Claimant per Month (1950–1991)

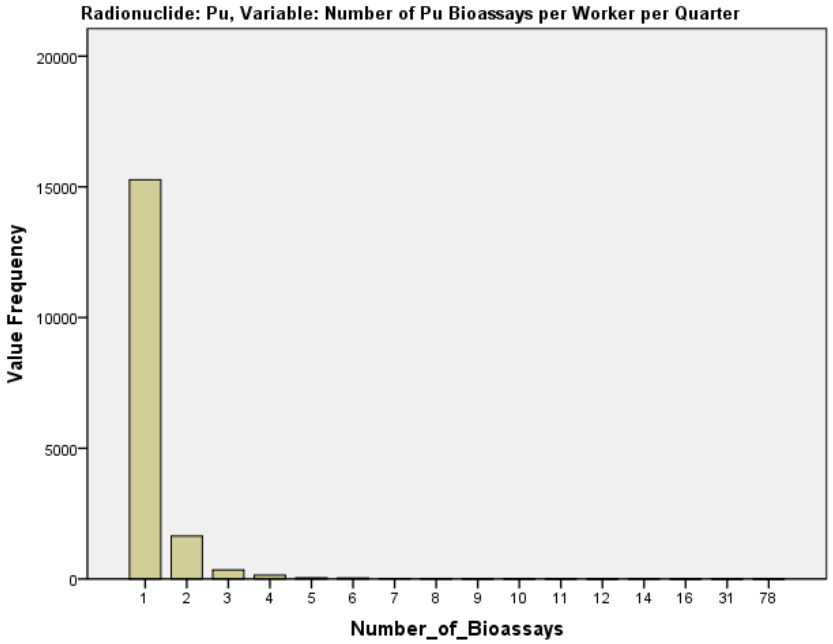

Figure A-2. Frequency Distribution of the Number of Plutonium Bioassay Results per Claimant per Quarter (1950–1991)

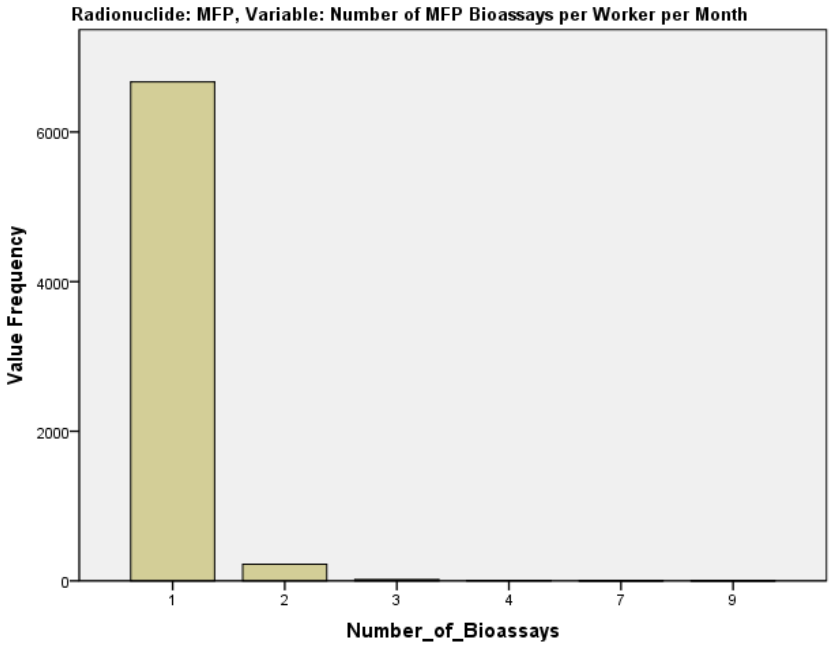

Figure A-3. Frequency Distribution of the Number of MFP Bioassay Results per Claimant per Month (1950–1991)

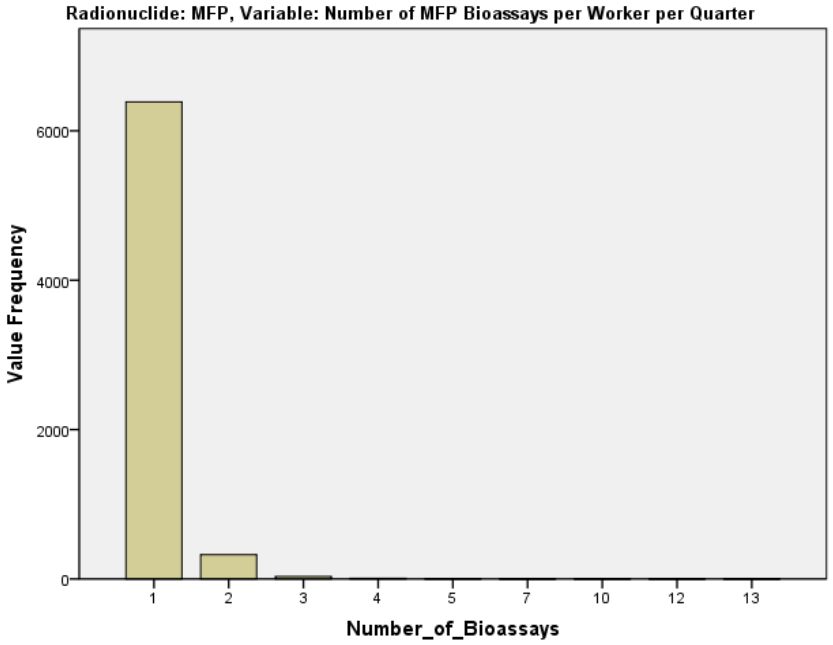

Figure A-4. Frequency Distribution of the Number of MFP Bioassay Results per Claimant per Quarter (1950–1991)

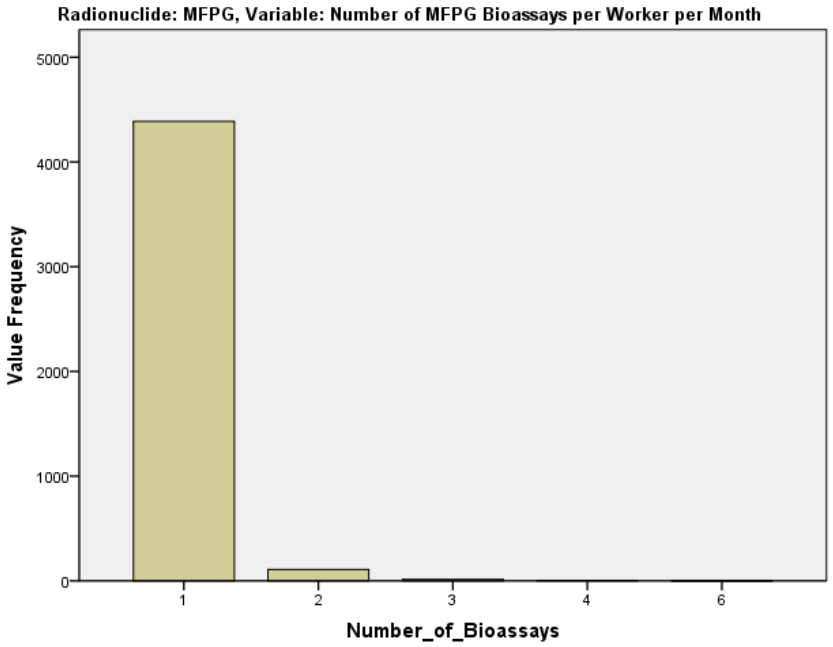

Figure A-5. Frequency Distribution of the Number of MFPG Bioassay Results per Claimant per Month (1950–1991)

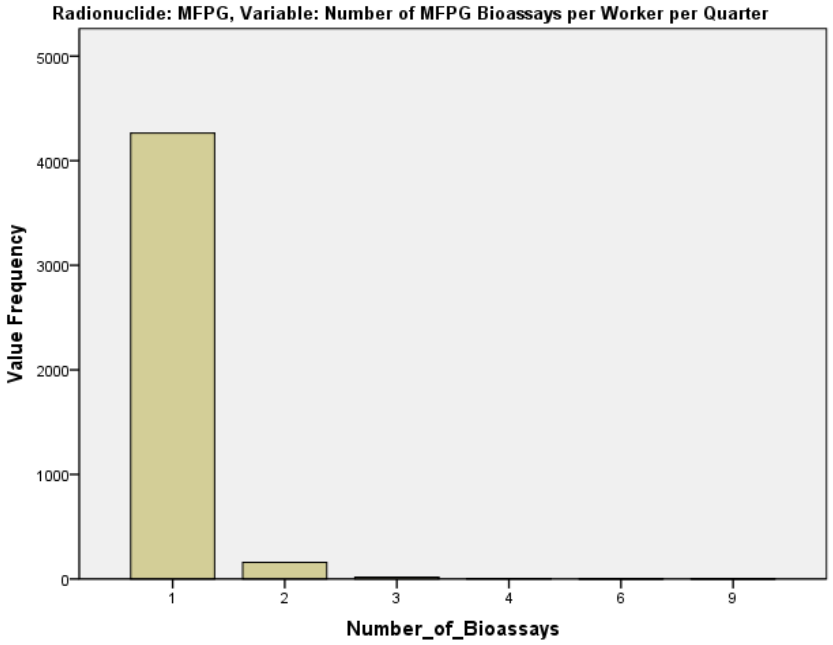

Figure A-6. Frequency Distribution of the Number of MFPG Bioassay Results per Claimant per Quarter (1950–1991)

Figure A-7. Frequency Distribution of the Number of Americium Bioassay Results per Claimant per Month (1950–1991)

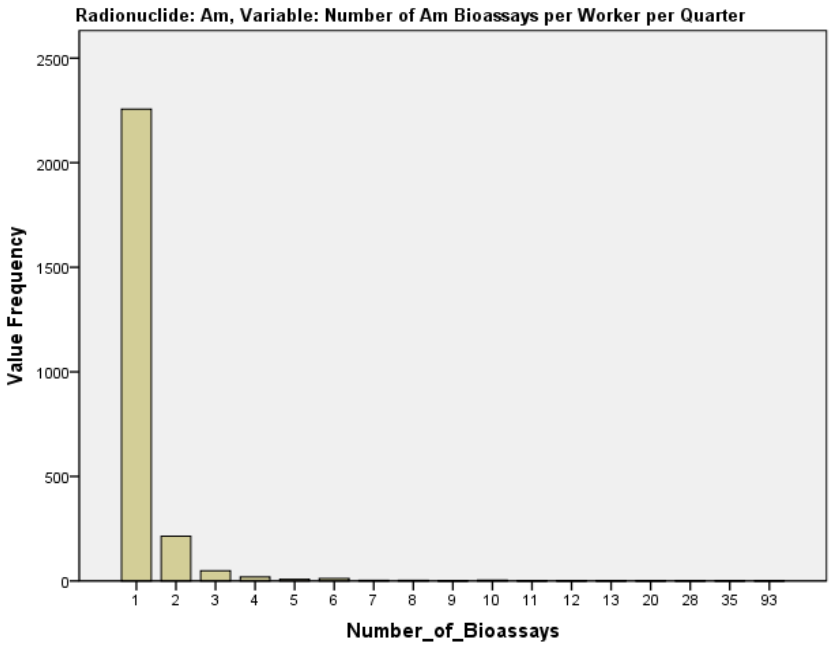

Figure A-8. Frequency Distribution of the Number of Americium Bioassay Results per Claimant per Quarter (1950–1991)

Figure A-9. Frequency Distribution of the Number of Uranium Bioassay Results per Claimant per Month (1950–1991)

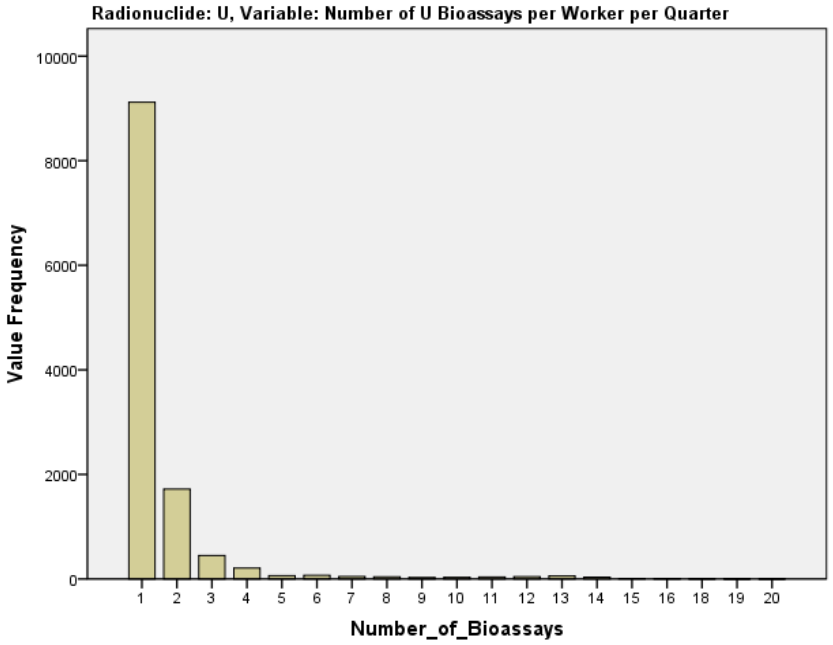

Figure A-10. Frequency Distribution of the Number of Uranium Bioassay Results per Claimant per Quarter (1950–1991)

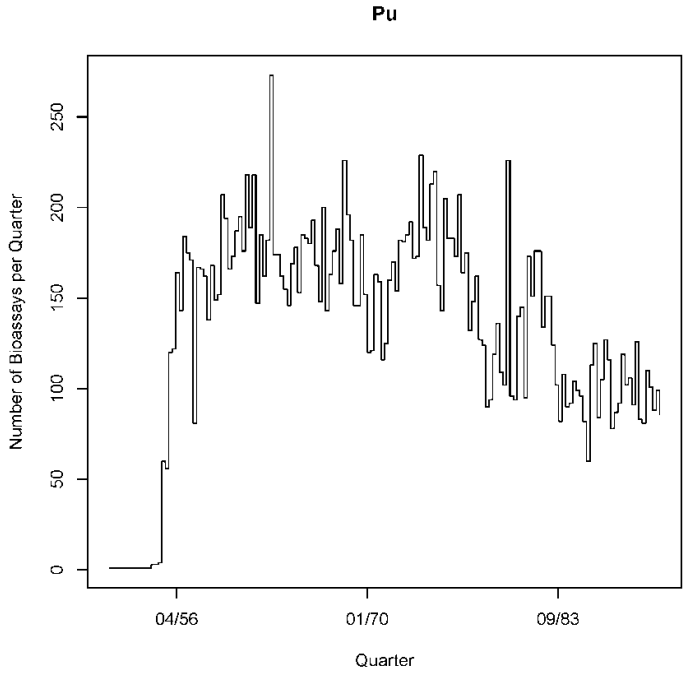

Figure A-11. Number of Plutonium Bioassay Results per Quarter (1950–1991)

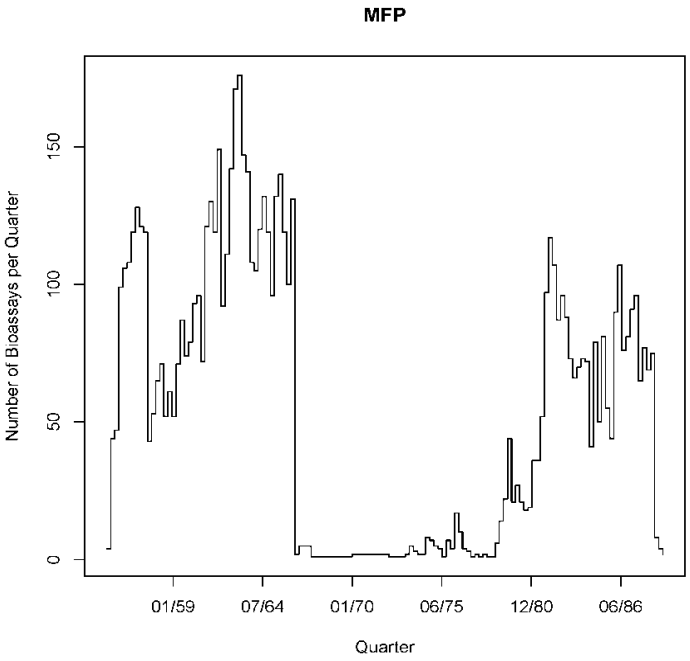

Figure A-12. Number of MFP Bioassay Results per Quarter (1950–1991)

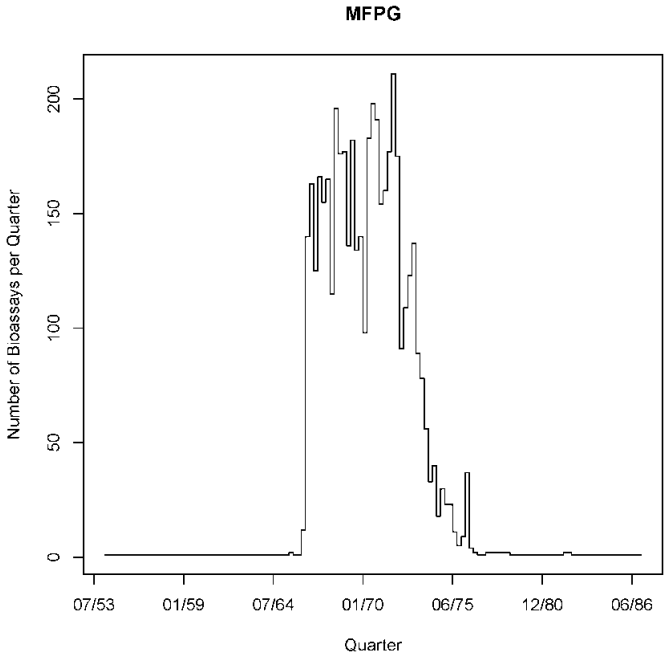

Figure A-13. Number of MFPG Bioassay Results per Quarter (1950–1991)

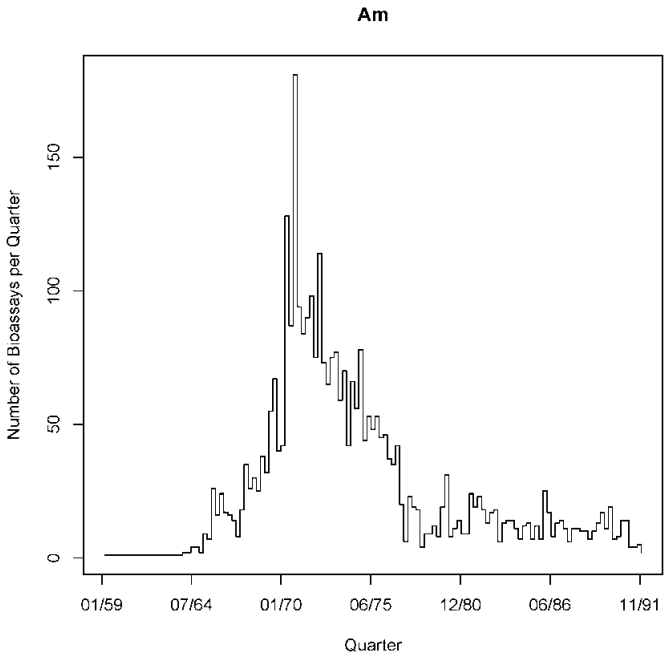

Figure A-14. Number of Americium Bioassay Results per Quarter (1950–1991)

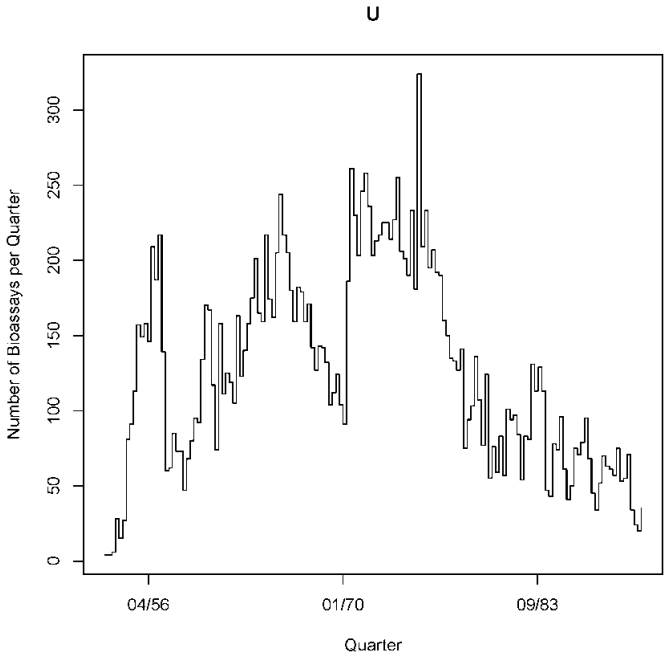

Figure A-15. Number of Uranium Bioassay Results per Quarter (1950–1991)

APPENDIX B: OPOS-RELATED FINDINGS IN SC&A REVIEW OF PROPOSED THORIUM AND NEPTUNIUM COWORKER MODELS AT SRS

B.1 OPOS-SPECIFIC FINDINGS FOR SRS THORIUM COWORKER MODEL:

SC&A's review of the SRS Evaluation Report Addendum 3 contained an overall finding on the OPOS approach that NIOSH is proposing to use for its coworker model for thorium and for the trivalent actinides: Am, Cm, and Cf. This overall finding is shown below.

Overall Finding on the One Person-One Sample Database and Model for Thorium and the Trivalent Actinides: The OPOS approach that NIOSH has adopted for trivalent actinides and thorium contains many fundamental problems, including large numbers of negative OPOS values in several years. In addition, many of the raw data records also appear to be unreliable, since widely different measurement results were obtained from different discs prepared from the same bioassay sample. As they stand, the problems with the data, the coworker model, and the OPOS results are severe enough to make their use for dose reconstruction scientifically unacceptable.

Additionally, SC&A had the following specific findings with regard to the use of the OPOS approach in the SRS thorium coworker model.

Finding 5: SC&A has concluded that NIOSH's method for comparing the measurements of two sets of workers requires that the monitoring protocols of the two sets of workers were the same. NIOSH has stated that the protocol for CTW bioassays was different. As a result, the method used by NIOSH to compare CTW and NCW Am/Cm/Cf data does not meet the requirements for a valid comparison of the two bioassay datasets for the 1972–1989 period.

Findings 6, 8: NIOSH's coworker model for thorium is based on its conclusion that CTW and NCW bioassay samples are drawn from the same distribution. The number of CTW data points is less than 30 in each aggregated period during 1984–1989. This is less than the minimum number required for a valid comparison between CTWs and NCWs. While NIOSH has not provided disaggregated data for 1981 and 1982, the number of CTW data points for 1982 is less than 30. Hence, the data for 1982 are also insufficient for a CTW-NCW distribution comparison. Therefore, NIOSH's conclusion that CTW and NCW sample distributions are the same is not valid for this period.

A corollary of Finding 5 above is that NIOSH's coworker model, which combines NCW and CTW data, is based on an invalid comparison and therefore is not suitable for estimating CTW thorium doses for the 1972–1989 period.

Finding 10: Aggregating data over more than 1 year without reference to underlying processes and other data is not justifiable. NIOSH should provide a technical rationale for treating 1981–1982 and 1987–1989 differently than other years. Aggregation over more than 1 year to increase the number of data points is not a suitable technical rationale. If no sound basis can be provided for aggregating data over more than 1 year, NIOSH should do annual aggregating for calculating OPOS values. This is important for evaluating NIOSH's conclusion that CTW and NCW data

are drawn from the same distribution. Furthermore, aggregation over multiple years rather than a single year to estimate an OPOS value increases the risk that the result would represent a mix of thorium exposure and Am/Cm/Cf exposure, rendering it scientifically questionable.

Finding 11: NIOSH has not demonstrated that the number of CTW samples is sufficient to simultaneously maintain low levels of Type 1 and Type 2 errors (for instance, less than 5% for Type 1 errors and less than 15% for Type 2 errors), even in the years when CTWs have more than 30 samples. SC&A's analysis indicates that when the geometric standard deviation (GSD) is much larger than the ratio of CTW to NCW geometric means (GMs), the rate of Type 2 errors will tend to be high. Type 2 errors occur when the null hypothesis (distributions are the same) is incorrectly accepted.

Findings 14, 19: NIOSH's approach to using data well below the MDA, including negative numbers and zeros to calculate OPOS values, can sometimes yield scientifically meaningless results such as negative OPOS values, implying negative intakes. The problem of negative OPOS results is especially prevalent in the 1983–1989 period.

Many reported OPOS values that are above the detection limit are actually the average of negative and positive normalized disc results, or are the average of results with large differences among the different discs derived from the same urine sample. Such average results no longer retain an unambiguous connection to the intake of the worker, do not represent excretion rates of workers, and therefore should not be used to calculate intake rates.

Finding 16: SC&A is concerned that some reported results in the logbooks that are above the MDA are averages of results that are both well below and well above the MDA. This is much better than the NIOSH OPOS procedure when even below MDA results are used at face value, but it is still a concern since such practices vitiate the connection between the raw data and the workers' intake experience in the real world.

Finding 20: Many reported OPOS results below the detection limit are the average of normalized disc results that have a large variation between them. This indicates that the resultant average of disc results is highly uncertain. Such average results do not have an unambiguous connection to the intake of workers, do not represent excretion rates, and should not be used to calculate intake rates.

B.2 OPOS-SPECIFIC FINDINGS FOR SRS NEPTUNIUM COWORKER MODEL

SC&A's review of the proposed methods for neptunium dose estimation produced the following OPOS specific findings.

Findings 1 and 2: SC&A has concluded that NIOSH's method for comparing the measurements of two sets of workers requires the monitoring protocols of the two sets of workers to be the same. NIOSH has not established that there were protocols for either whole-body count (WBC) during the 1972–1990 period, nor bioassay sampling in the 1991–2007 period, of either non-construction trades workers (NCWs) or construction trades workers (CTWs), and if so, whether they were comparable. Establishing that equivalence is a necessary (though not sufficient)

condition for the application of a coworker model based on "all-worker" monitoring results to CTW's. It appears unlikely that either group (as a whole) was routinely counted, except for "fast scans" in the mid-to late 1980s, which are not relevant to neptunium dose reconstruction.

Findings 3, 4, and 7: NIOSH has not demonstrated that 30 samples in each comparison group (CTWs and NCWs) would be sufficient to simultaneously maintain low levels (for instance, less than 10% for each) of Type 1 and Type 2 errors in determining whether CTW and NCW sample distributions are the same (or not).

However, the issue is moot in the case of neptunium only because the minimum number of 30 samples for the comparison to be valid is not available for CTWs in any year in the 1972–1990 period. That is, NIOSH's method for concluding that CTW measurements were drawn from the same distribution as those for NCWs does not meet the minimum technical requirements for 30 samples from each of the groups being compared for the 1972–1990 period. There were fewer than 30 total CTW samples in each year in the 1972–1990 period for which NIOSH reported data; the number of above minimum detectable activity (MDA) results ranged from just 6 to a high of 17 for the same period. As a result, the coworker model is based on a comparison that does not satisfy the minimum data adequacy criteria.

Finding 8: A statistically valid analysis comparing neptunium exposure potential of subgroups of CTWs or of relative CTW exposure potential in various SRS areas is not possible at the present time for the 1991–2007 period, since the data for that period have not been separated into CTW and NCW categories.

APPENDIX C: OPOS-RELATED FINDINGS IN THE REVIEW OF OTIB-0078: INTERNAL DOSE COWORKER DATA FEMP, REV. 2

C.1 Background and Introduction

ORAUT OTIB-0078, Internal Dose Coworker Data FEMP, Rev. 2 (ORAUT 2012a) provides uranium coworker intakes at Fernald. SC&A understands that, at the time of this report, ORAUT 2012a is the only coworker model other than those pertaining to SRS that employs the OPOS methodology. In addition, the previous version of the OTIB-0078, Revision 1 (ORAUT 2010) uses the previous methodology (referred to herein as the "pooled" approach) to compute the intake rates of the unmonitored worker.

Given the large amount of uranium urine bioassay data available at FEMP that spans more than five decades and covers approximately 90 percent of employees and the relatively high sampling frequency, the Fernald dataset provides a unique opportunity to evaluate the effects of the application of OPOS as compared to the previous methodology and will help elucidate the extent of data dominance and correlation. In addition, SC&A has evaluated whether some of the implementation problems encountered in the SRS coworker models are isolated or potentially systemic.

As can be seen in Table 2-3 of ORAUT 2012a, the 50th and 84th percentile uranium excretion rates based on annual OPOS results were calculated based on annual intervals from 1952 through 2006. These annual 50th and 84th percentile OPOS-derived uranium excretion rates were used to determine appropriate 'intake regimes' (periods where bioassay results are sufficiently similar in magnitude to conclude that intake rates are likely the same). The number of OPOS values available per year is shown in Figure C-1. The figure shows that the number of available OPOS results ranged from a low of 71 (1952) to a high of 3,166 (1993); on average, there were approximately 1,600 results per year.

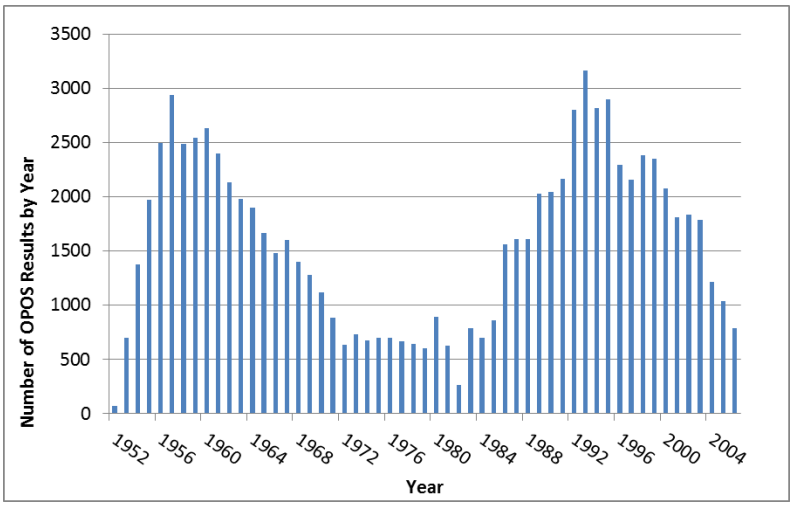

Figure C-1: Number of Annual OPOS Results by Year from OTIB-0078 Rev. 2

The corresponding 50th percentile urine excretion rates result for each year is shown in Figure C-2. The highest observed 50th percentile results occurred in the mid-1950s with the highest 50th percentile result in 1955 (46.62 µg/d). This likely corresponds to the site beginning to increase overall uranium operations while health and safety operations were still evolving. Beginning in the early 1960s, 50th percentile results generally declined until there was another brief spike beginning in 1991, which continued through 1993. However, the 1991–1993 values may be a result of the monitoring practices and of changes in reporting practices and do not necessarily reflect an increase in exposure potential.

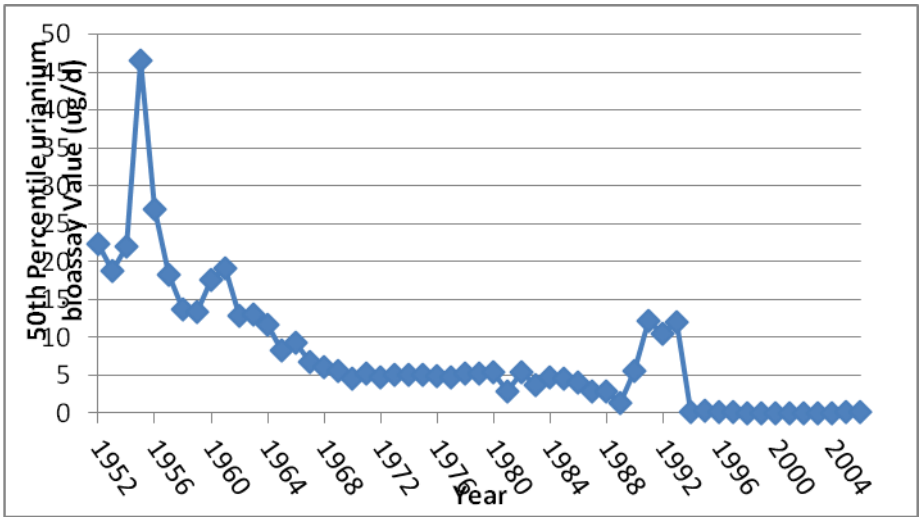

Figure C-2: 50th Percentile Bioassay Result by Year (µg/d)

The chosen intake regimes and calculated intake rates based on these 50th percentile uranium urinary excretion rate values in units of micrograms per day (µg/d) are shown in Table C-1 for each solubility type. As can be seen in columns 2 and 3 of the table, intake regimes varied in length from a single year for the two highest observed annual 50th percentile values to as much as 27 years (1967–1993). Also, the magnitude of 50th percentile urine excretion rates values in a given regime vary by as much as a factor of 10, as is the case in the 1994–2006 period.

Table C-1: Characterization of Chosen Intake Regimes in OTIB-0078 Rev. 2

Dates of Intake Regime*	Number of Years in Regime	Range of Annual 50th Percentile Urine Excretion Rates in Regime (µg/d)	50th Percentile Intake Rate for Type F Uranium (µg/d)	50th Percentile Intake Rate for Type M Uranium (µg/d)	50th Percentile Intake Rate for Type S Uranium (µg/d)
01/01/1952–12/31/1954	3	18.82–22.3	77.6	334	7,393
01/01/1955–12/31/1955	1	46.62	171.9	770.4	26,230
01/01/1956–12/31/1956	1	26.85	97.17	339.5	15,080
01/01/1957–12/31/1961	5	13.34–19.05	58.81	235.7	4,681
01/01/1962–12/31/1966	5	8.21–12.99	39.06	156.8	2,999
01/01/1967–12/31/1993	27	1.28–6.78	16.14	65.17	799.1
01/01/1994–12/31/2006	13	0.021–0.279	0.365	1.487	17.84

*ORAUT 2012a, Tables 5-1 to 5-3 often break up the shown intake regimes into smaller components to reflect temporal changes in uranium enrichment and hence changes in calculated intake values provided in activity per day.

During various periods of time at Fernald, some workers collected a high number of urine samples, many of them taken on the same day. Figure C-3 shows that the maximum number of samples taken from a single worker per quarter could range as high as nearly 200 samples.

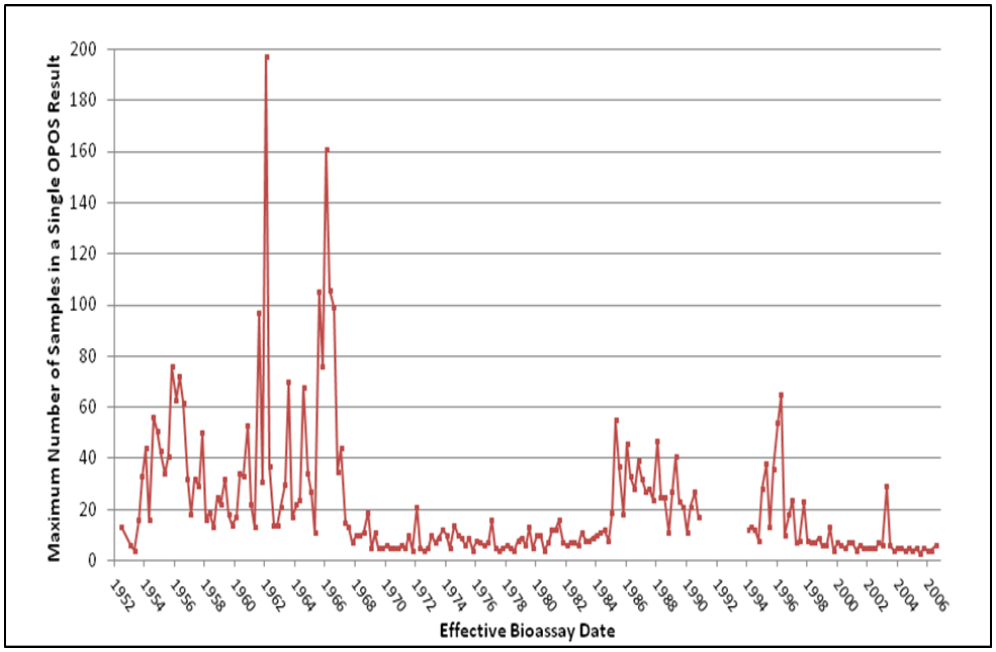

Figure C-3: Maximum Number of Samples per Quarter Taken from a Single Worker

SC&A notes that it had no access to the raw calculation of OPOS for each individual worker in ORAUT 2012a. However, SC&A had access to the raw urinalysis database for Fernald, as well as the dataset used in ORAUT 2010. In the latter case, all the lognormal distributions are posted on the O-Drive at the following location [O:\Coworker Data\Approved Files\Internal Dose data\Fernald\Uranium in urine]. SC&A examined some lognormal distributions in this directory for ORAUT 2010 and determined that the results from the same day appear to be used independently, which is theoretically not correct. In the case of calculating the OPOS either for the limited time period of special sampling or annually, as calculated by NIOSH in ORAUT 2012a, all results from the same day should be combined to produce one value prior to calculating the OPOS result for a given period of time. The ideal situation would be to know the volume of each of the samples taken on the same day. Then, for each day, the results in activity per liter would be multiplied by the sample volume and added to give the excretion rate for each day. If there is no such information, the average activity per liter may be calculated for each single day.

To ascertain whether NIOSH has used the same-day results as independent samples or has combined the same day results and averaged them for ORAUT 2012a. SC&A attempted to replicate the 50th and the 84th percentile excretion rates from ORAUT 2012a using both adjusted daily samples (multiple samples in a single day are averaged first) or unadjusted daily samples (multiple daily samples used independently). Since the calculation files used in ORAUT 2012a were not currently available, SC&A used the raw urinalysis values in the HIS_20 database to try and estimate the annual OPOS results. The results are shown in Table C-2. As seen in the table,

in many cases SC&A could not determine how NIOSH used the same day samples since calculation of the annual 50th and 84th percentile excretion result was nearly identical using either method for most years.

Table C-2: Comparison of OPOS 50th and 84th Percentile Excretion Rates from ORAUT 2012a with SC&A Calculations, using the Average of Same-Day Samples or using the Results as Independent Samples

			50th Excretion (µg/d)			84th Excretion (µg/d)		
Year	**# OPOS**	**% Raw Results Occurring For the same Worker on the Same Day**	**OTIB-0078 Rev. 2**	**SC&A Daily results Averaged**	**SC&A Daily results used independently**	**OTIB-0078 Rev. 2**	**SC&A (Daily results Averaged**	**SC&A Daily results used independently)**
1952	71	6.05%	22.30	21.79	22.25	69.16	50.18	66.43
1953	701	5.73%	18.82	24.44	18.81	55.73	65.45	55.39
1954	1376	4.09%	22.08	22.10	22.08	60.42	60.40	60.24
1955	1973	6.52%	46.62	46.54	46.61	116.98	116.61	116.75
1956	2497	2.06%	26.85	26.82	26.84	68.80	68.51	68.68
1957	2937	5.63%	18.27	18.26	18.27	50.25	50.12	50.16
1958	2485	2.95%	13.67	13.65	13.67	34.01	33.88	33.95
1959	2540	0.63%	13.34	13.33	13.34	28.45	28.37	28.41
1960	2630	1.99%	17.61	17.58	17.61	32.79	32.67	32.80
1961	2395	6.54%	19.05	19.05	19.05	33.62	33.57	33.59
1962	2131	6.00%	12.82	12.79	12.81	25.14	24.98	25.12
1963	1983	6.09%	12.99	12.98	13.01	25.70	25.63	25.67
1964	1900	2.44%	11.74	11.72	11.75	24.50	24.31	24.46
1965	1663	11.67%	8.21	7.81	7.91	17.65	17.80	18.06
1966	1484	16.18%	9.36	9.26	9.25	23.00	21.91	24.42
1967	1602	1.53%	6.78	6.78	6.78	15.44	15.40	15.41
1968	1398	2.77%	6.07	6.06	6.06	13.85	13.76	13.79
1969	1281	0.58%	5.53	5.53	5.53	12.69	12.92	12.93
1970	1119	1.13%	4.51	4.51	4.51	9.95	9.91	9.92
1971	881	0.05%	5.30	5.33	5.30	11.75	12.06	11.70
1972	634	0.24%	4.65	4.65	4.65	12.14	12.05	12.07
1973	735	0.16%	5.03	5.03	5.03	11.83	11.75	11.77
1974	678	0.04%	4.99	4.99	4.98	11.42	11.36	11.36
1975	697	0.54%	5.00	5.00	5.00	10.83	10.78	10.78
1976	697	0.30%	4.90	4.90	4.90	10.47	10.42	10.43
1977	664	0.06%	4.80	4.80	4.80	10.02	9.98	9.98
1978	644	0.06%	5.17	5.17	5.17	10.32	10.27	10.27
1979	599	0.48%	5.26	5.26	5.26	10.52	10.49	10.47
1980	893	0.59%	5.40	5.40	5.39	12.23	12.18	12.12
1981	623	0.70%	2.89	2.89	2.89	7.13	7.10	7.09
1982	262	0.75%	5.42	5.42	5.42	10.44	10.45	10.44
1983	785	0.97%	3.77	3.77	3.74	8.85	8.81	8.82
1984	696	1.34%	4.71	4.70	4.70	10.93	10.85	10.89
1985	858	5.15%	4.63	4.61	4.63	9.07	8.95	9.04
1986	1565	6.01%	4.07	4.59	4.59	6.27	6.20	6.22
1987	1611	2.50%	2.82	4.51	4.52	5.48	5.67	5.70
1988	1609	2.31%	2.92	4.25	4.25	4.75	5.12	5.13
1989	2029	1.60%	1.28	6.86	6.85	2.98	7.27	7.27
1990	2044	0.36%	5.51	8.73	11.99	7.21	10.62	12.55
1991	Not Used by NIOSH		–	–	–	–	–	–
1992	Not Used by NIOSH		–	–	–	–	–	–
1993	Not Used by NIOSH		–	–	–	–	–	–
1994	2817	0.17%	0.20	1.14	1.12	0.43	1.19	1.13

Table C-2: Comparison of OPOS 50th and 84th Percentile Excretion Rates from ORAUT 2012a with SC&A Calculations, using the Average of Same-Day Samples or using the Results as Independent Samples

			50th Excretion (µg/d)			84th Excretion (µg/d)		
Year	**# OPOS**	**% Raw Results Occurring For the same Worker on the Same Day**	**OTIB-0078 Rev. 2**	**SC&A Daily results Averaged**	**SC&A Daily results used independently**	**OTIB-0078 Rev. 2**	**SC&A (Daily results Averaged**	**SC&A Daily results used independently)**
1995	2901	0.34%	0.28	1.12	1.12	0.54	1.14	1.14
1996	2298	2.18%	0.11	1.13	1.13	0.34	1.15	1.15
1997	2159	0.39%	0.232	0.268	0.268	0.722	0.870	0.871
1998	2382	0.25%	0.021	0.021	0.021	0.071	0.070	0.070
1999	2351	0.07%	0.025	0.025	0.025	0.070	0.070	0.070
2000	2076	0.13%	0.034	0.034	0.034	0.097	0.097	0.097
2001	1809	0.21%	0.065	0.065	0.065	0.145	0.144	0.144
2002	1835	0.25%	0.062	0.062	0.062	0.130	0.130	0.130
2003	1787	0.16%	0.056	0.056	0.056	0.134	0.134	0.134
2004	1217	0.21%	0.058	0.058	0.058	0.115	0.115	0.115
2005	1040	0.21%	0.091	0.091	0.091	0.167	0.167	0.167
2006	790	0.04%	0.082	0.082	0.082	0.154	0.154	0.153

SC&A could not recreate the results found in ORAUT 2012a for the years 1987–1990 and 1994–1997 (highlighted) using either method (same-day samples used independently or averaged).

The comparisons in Table C-2 demonstrate that while treating same day as independent samples is scientifically incorrect, doing so had no apparent impact on the derived uranium excretion rate results at either the 50th or 95th percentiles.

C.2 Comparison of the Intake Rates Assigned to the Unmonitored Worker in OTIB-0078 Rev. 1 (2010) and in OTIB-0078 Rev. 2 (2012a) – Effect of Addition of Samples from Special Programs, Characterized by a Large Number of Samples from the Same Worker (referred to by NIOSH as "Data Dominance")

According to NIOSH, one of the principal purposes of the OPOS methodology was to take into account multiple samples from the same individual in a given period "which creates a related problem if an individual is involved in an incident and has more (potentially many more) bioassay results than other workers" (RPRT-0053, 2012b). In theory, including such results could distort the distribution of results used to derive intakes for unmonitored workers. NIOSH has derived coworker intake rates to assign doses to the unmonitored worker for many sites. In some of those sites, those multiple samples from the same worker are discarded.

In ORAUT 2010, NIOSH did not consider samples with the Code 50 that were collected as part of a special study. In ORAUT 2012a Code 50 results were used in the OPOS calculation for each individual worker. It is noteworthy that both versions of OTIB-0078 take into account results labeled 40 and 49, which are incident-related samples. SC&A notes that it is difficult to determine the impact of "data dominance" in ORAUT 2012a without first looking at the raw data that were used to generate the OPOS values for each worker – an "apples to apples" comparison; at the time of this report, SC&A did not know which data were used in the OPOS calculations.

For SC&A to ascertain whether data dominance truly exists in the Rev. 2 dataset, NIOSH will need to identify the raw data used in the Rev. 2 OPOS calculations.

It is worth restating that the main objective of the analysis of monitored coworker information is to calculate and assign occupational internal doses to employees for whom there are no or insufficient bioassay monitoring records. This is accomplished through the assignment of intake rates to the unmonitored worker. SC&A compared the 50th and 95th percentiles intake rates from ORAUT 2010 and ORAUT 2012a, to determine if the inclusion of Code 50 data had influenced the intake rates in ORAUT 2012a.

It must be acknowledged that a comparison of the ORAUT 2010 and ORAUT 2012a intake rate models involves some confounding factors. First and most obvious, the two models rely on different datasets; that is, ORAUT 2010 does not consider Code 50 data and ORAUT 2012a does use Code 50 data. Second, the calculated intake rates will be affected by the intake regime (period over which excretion rates are grouped together to obtain a single intake rate). ORAUT 2010 and ORAUT 2012a do not consider the same intake regimes to calculate intake rate from the available excretion data. For example, it can be seen in Table C-3 that ORAUT 2012a calculated a single intake rate at the 50th percentile for the 1957–1961 intake regime (column 2), while ORAUT 2010 calculated a single intake rate for just the 1957–1958 period (column 4).

The third variable affecting this intake comparison is the specific date at which the intake is evaluated. ORAUT 2012a calculated the intake rate based on an annual excretion result evaluated at the midpoint of each year in the intake regime (July 1st of the given year), while ORAUT 2010 calculated the intake rate based on the quarterly excretion values evaluated at the midpoint of each quarter (2/15, 5/15, 8/15, and 10/15 of a given year). Nonetheless, with these caveats noted, a comparison of the calculated intake rates from ORAUT 2012a and ORAUT 2010 can still provide some insight into the potential effect of data dominance that could result from inclusion of Type 50 sample results. Tables C-3 and C-4 below help to elucidate the impact of data dominance on the respective models.

Table C-3 below compares the intake rates for Type S uranium from ORAUT 2010 Revision 1 with the intake rates derived from ORAUT 2012a. For each period of time, the number of samples with labels 50 is given. Some samples with label 50 or 5c are repeated samples taken on the same day, from the same worker.

Table C-3. Comparison of Intake Rates for Type S Uranium as Assigned in OTIB-0078 Rev. 1 (2010) and OTIB-0078 Rev. 2 (2012a)

Intake Dates	50th percentile Intakes from Rev. 2 (2012) (µg/d)	95th percentile Intakes from Rev. 2 (2012) (µg/d)	50th percentile Intakes from Rev. 1 (2010) (µg/d)	95th percentile Intakes from Rev. 1 (2010) (µg/d)	Number of Samples with Labels 50
1/1/1952–12/31/1953	7,393	45,049	8,197	62,559	0
1/1/1954–12/31/1954	7,393	45,049	15,042	91,658	0
1/1/1955–12/31/1955	26,230	159,832	15,042	91,658	1
1/1/1956–12/31/1956	15,080	91,889	15,042	91,658	0
1/1/1957–12/31/1957	4,681	28,524	7,140	68,407	17
1/1/1958–12/31/1958	4,681	28,524	7,140	43,507	1,767

Table C-3. Comparison of Intake Rates for Type S Uranium as Assigned in OTIB-0078 Rev. 1 (2010) and OTIB-0078 Rev. 2 (2012a)

Intake Dates	50th percentile Intakes from Rev. 2 (2012) (µg/d)	95th percentile Intakes from Rev. 2 (2012) (µg/d)	50th percentile Intakes from Rev. 1 (2010) (µg/d)	95th percentile Intakes from Rev. 1 (2010) (µg/d)	Number of Samples with Labels 50
1/1/1959–12/31/1960	4,681	28,254	7,772	47,358	4,573
1/1/1961–12/31/1961	4,681	28.254	3,628	22,107	664
1/1/1962–12/31/1965	2,999	18,274	3,628	22,107	2,843
1/1/1966–12/31/1966	2,999	18,274	3,628	22,107	228
1/1/1967–12/31/1980	799.1	4,869	3,628	22,107	996
1/1/1981–12/31/1986	799.1	4,869	1,252	7,629	1,497
1/1/1987–12/31/1990	799.1	4,869	361	2,396	2,781
1/1/1991–12/31/1993	799.1*	4,869*	361**	2,396**	
1/1/1994–6/30/1997	17.84	108,7	1.34	164,2	2,088
7/1/1997–12/31/1999	17.84	108,7	7.29	54,8	242
1/1/2000–12/31/2006	17.84	108,7	14.0	85,3	852

* For 1991–1993, the results were excluded

** For 1991 through 1993, fewer than 10 results per year were above the minimum detectable activity, which resulted in an insufficient quantity of data to generate a meaningful intake model. Intake rates from the previous interval were assumed.

Table C-4 below shows the number of samples labeled special samples (code 50 or 5C) and the number of incident-related samples (code 40 or 49) for 1952 to 1965. For this example, the intake rates for the respective models are compared in light of the number of incident-related and special samples. It is evident from the comparison that differences in the intake rates for the respective models cannot always be explained by "data dominance."

Table C-4. Number of OPOS Results as Compared to the Number of Special Samples and Incident-related Samples per Year from 1952 to 2006

Year	# OPOS	Special Samples (Type 50s)	Incident Samples (Type 40 or 49)
1952	71	0	0
1953	701	0	0
1954	1,376	0	2
1955	1,973	1	44
1956	2,497	0	13
1957	2,937	17	25
1958	2,485	1767	2161
1959	2,540	1,731	1,080
1960	2,630	2,842	1,827
1961	2,395	664	535
1962	2,131	356	992
1963	1,983	2,053	893
1964	1,900	133	439
1965	1,663	301	335
1966	1,484	228	946
1967	1,602	43	214
1968	1,398	434	78
1969	1,281	43	50
1970	1,119	0	37
1971	881	7	16

Table C-4. Number of OPOS Results as Compared to the Number of Special Samples and Incident-related Samples per Year from 1952 to 2006

Year	# OPOS	Special Samples (Type 50s)	Incident Samples (Type 40 or 49)
1972	634	269	18
1973	735	16	114
1974	678	29	161
1975	697	6	71
1976	697	21	44
1977	664	35	22
1978	644	8	23
1979	599	47	14
1980	893	38	33
1981	623	71	106
1982	262	24	123
1983	785	21	239
1984	696	87	359
1985	858	108	3,317
1986	1,565	1,186	2,052
1987	1,611	1,094	1,447
1988	1,609	731	935
1989	2,029	939	600
1990	2,044	17	351
1991	Not Used by NIOSH		
1992	Not Used by NIOSH		
1993	Not Used by NIOSH		
1994	2,817	450	182
1995	2,901	678	178
1996	2,298	853	146
1997	2,159	154	133
1998	2,382	122	184
1999	2,351	73	81
2000	2,076	120	66
2001	1,809	178	105
2002	1,835	102	79
2003	1,787	224	99
2004	1,217	217	49
2005	1,040	11	41
2006	790	0	27

No samples coded 50 were taken in 1952–1956, with the exception of 1 sample in 1955. In 1954 there were no samples labeled 50 and only two samples labeled 40. Both the 50th percentile and the 95th percentile intake rates in ORAUT 2010 were two times higher than in ORAUT 2012a. There were no multiple samples to distort results and yet ORAUT 2010 assigned higher intake rates to the unmonitored worker. Thus, from 1953 to 1954, the shift of which version of OTIB-0078 assigns the higher intake rate was not due to the collection of multiple samples because of incidents or special samples.

In 1955, there were 1,973 OPOS results and only 44 incident-related samples. There were no samples labeled 50. Both the 50th percentile and the 95th percentile intake rates in ORAUT 2012a were 1.7 times higher than in ORAUT 2010.

In 1957, there were 2,947 OPOS results and only 17 special samples and 25 incident-related samples. The number of incident samples could not have distorted the 50th percentile intake rates of OTIB-2010. Yet the 50th percentile intake rate from ORAUT 2010 was 1.5 higher than in ORAUT 2012a.

In 1963, there were 1,983 OPOS results, 2,053 special samples and 893 incident samples. In 1964, there were 1,900 OPOS results and the number of special samples dropped to 133 samples and the incident-related samples to 439. Yet, the 50th and 95th percentile intake rates assigned do not change from 1963 to 1964, in both ORAUT 2010 and ORAUT 2012a.

ORAUT 2010 assigns slightly higher 50th and 95th percentile intake rates than ORAUT 2012a. There was no influence resulting from the decreased number of special samples from 1963 to 1964.

Until 1986, with the exception of 1955, 1956 and 1961, the intake rates derived in ORAUT 2010 were higher than the intake rates derived in ORAUT 2012a. The opposite occurs after 1987.

In 1987, there were 1,611 OPOS results and 1,094 special samples. In 1990, there were 2,044 OPOS results and only 17 special samples. Yet, the 50th and 95th percentile intake rates assigned do not change from 1987 to 1990, in both ORAUT 2010 and ORAUT 2012a.

ORAUT 2012a assigns two times higher 50th and 95th percentile intake rates than ORAUT 2010. There was no influence resulting from the decrease on the number of special samples from 1987 to 1990.

The inclusion of samples labeled 50 in the Rev. 2 model did not influence the assignment of intake rates. Furthermore, the comparison of assigned intake rates in ORAUT 2010 and ORAUT 2012a for the periods 1952 to 1957 and 1962–1963 demonstrate that "data dominance" had no influence on which version of the OTIB assigned the higher intake rate.

Finding C-1: The comparison of the 50th and 95th percentiles intake rates derived in Revision 1 (2010) and Revision 2 (2012a) of OTIB-0078 has shown that the addition of samples from special programs in the Revision 2 model, characterized by a large number of samples from the same worker, did not increase or influence the intake rates for the unmonitored worker when compared to the Revision 1 model. "Data dominance" from incident or special monitoring programs had no influence on which version of the OTIB-0078, ORAUT 2010 or ORAUT 2012a, assigned the higher intake rate.

C.3 Correlation

NIOSH's second principal argument for data reduction using OPOS is to avoid the effects of "correlation." It is not clear whether this statement refers to the possible serial correlation that may occur in samples over time for a given worker, or perhaps to intra-cluster correlation.

In this context, both versions of the OTIB take the same approach in relation to the influence of results from one time period to the other.

ORAUT 2012a states:

> *Because the uranium isotopes at FEMP have long radiological half-lives and the material is retained in the body for long periods, excretion results are not independent. For example, an intake in the 1950s could contribute to urinary excretion in the 1980s and later. To avoid potential underestimation of intakes for people who worked at FEMP for relatively short periods, each chronic intake was fit independently using only the bioassay results from the single intake period for type S solubility. For types M and F solubility, this approach was used where it was determined that earlier intake rates significantly biased later intake rates, i.e., 1994 through 2006 was evaluated separately from earlier time periods. This method results in a potential overestimate of intakes for exposures that extend through multiple assumed intake periods.*

Similar sentence is seen in ORAUT 2010.

For data related to individual workers exposed in accidents or special work assignment(s), when a great number of samples are taken in sequential days during a definite time period, it is not clear why correlation problems end from one year to the other. OPOS is calculated for each calendar year, but the correlation among monitoring results related to the same worker does not follow the calendar year. It depends on the time of the exposure, the intake rate, the pattern of intake, the physical and chemical characteristic of the nuclide and the biokinetics of the nuclide in the human body. Thus, for Fernald, OPOS does not avoid the effects of correlation, which is present in OPOS and pooled data models.

Finding C-2: ORAUT 2012a applies the OPOS methodology to assign intake rates for the unmonitored worker. In the case of Fernald, the use of OPOS methodology has not resolved the dependence of monitoring results from different and sequential intake periods. ORAUT 2012a, explicitly exemplifies that for the 1994–2006 period, earlier intake rates significantly biased later intake rates, for all solubility types of uranium compounds.

C.4 Fernald Data as an Example to Illustrate the Influence of the Frequency of Monitoring when OPOS is Calculated as an Average of the Excretion Rates for the Whole Year

As pointed out in the main report, one of the problems of using OPOS for the whole year instead of applying it to a particular intake period is that the result will depend on the number of

monitoring results that the worker had in the year, in addition to the number of significant exposures; that is, **the OPOS result is directly dependent on the sampling frequency**. For example, consider a worker who was monitored for 11 months with results < MDA, had a high exposure in the last month of the year, and was monitored 5 times with results R1, R2, R3, R4 and R5, all higher than the MDA. The annual OPOS for this worker will be (11 × MDA +R1+R2+R3+R4+R5)/16. If this same worker was under a less stringent monitoring program and was only monitored in April with the same <MDA result, and then 5 times in December after he was exposed, with the same R1, R2, R3, R4 and R5, the annual OPOS of this worker would be (MDA +R1+R2+R3+R4+R5)/6; a higher result than the previous one though the intake was the same. The only difference is the frequency of monitoring before the exposure. Thus the individual OPOS result, as calculated on an annual basis, is influenced by the frequency of monitoring of the worker.

To illustrate this problem, SC&A reviewed the uranium urine excretion results from the HIS-20 database for the year of 1966; the year 1966 was chosen at random. It can be inferred from the 1966 sample data that an incident probably occurred in February of that year. Several samples from various workers were taken around the 14th to the 22nd of February. Several samples from the same worker were taken on the same day. The uranium concentration in urine indicated high excretion rates associated with samples from this incident.

SC&A calculated the 1966 annual OPOS for three workers that were monitored in the dates around the incident. Two of the workers had more than one sample taken on the day of the incident. As the volume of individual excretion rates were not given, SC&A calculated the average concentration for samples taken in the same day.

Worker A involved in the accident had high urine uranium concentration results. In 1966, he was only monitored on the 14th to the 17th of February. All his bioassay results had codes 40 (14th, 15th and 16th of February) or 49 (one of the samples taken on the 17th of February). ORAUT-TKBS-0017-5, 2004 defines the samples with code 40 as "Incident-Follow-up Sample-samples from employees involved in an event or circumstances which presents a potential for elevated exposure." The samples with code 49 are described as "an incident sample left at the end of the shift on the day of the incident."

The daily samples results were:

14th of February: 1,200 µg/L (code 40), 1,300 µg/L (code 40); average: 1,250 µg/L
15th of February: 56 µg/L (code 40), 79 µg/L (code 40); average: 67.5 µg/L
16th of February: 23 µg/L (code 40), 29 µg/L (code 40); average: 26 µg/L
17th of February: 7 µg/L (code 40), 12 µg/L (code 49); substituted by the MDA value of 14 µg/L on the MPM calculation

His MPM was calculated by SC&A as equal to 340 µg/L.

Worker B involved in the same accident had a higher frequency of monitoring, in February and later in the year, after the incident. Several results are given on the same day for this worker in February. His samples related to the incident have various codes. In addition to codes 40 and

49, the results have codes 5B, 5H and xx. ORAUT-TKBS-0017-5, 2004 defines the samples with code 5B as "off the job, overnight individual specimen" and 5H as "on-the-job individual sample collected in the work area." ORAUT-TKBS-0017-5, 2004 does not define samples with codes xx. Worker B had three samples taken during the year with code 30, defined in ORAUT-TKBS-0017-5, 2004 as "Routine Sample-samples from plant workers who are on a routine schedule for the purpose of insuring that airborne levels of uranium in the work place are being controlled with safe limits."

14th of February : 1,100 µg/L (code 40) , 700 µg/L (code xx), 360 µg/L (code 5B), 120 µg/L (code 5H), 200 µg/L (code 5H), 500 µg/L (code 5B), 580 µg/L (code xx), 190 µg/L (code 5H); average: 469 µg/L

15th of February: 160 µg/L (code 5B), 230 µg/L (code 5B), 410 µg/L (code 5B), 200 µg/L (code 40), 67 µg/L (code 40); average: 213 µg/L

16th of February: 28 µg/L (code 40), 16 µg/L (code 40); average: 22 µg/L

17th of February: 17 µg/L (code 40), 29 µg/L (code 49); average: 23 µg/L.

18th of February: 17 µg/L (code 40)

21st of February: 10 µg/L (code 40); substituted by the MDA value of 14 µg/L in the MPM calculation

24th of March (code 30): 7 µg/L; substituted by the MDA value of 14 µg/L in the MPM calculation

11th of July (code 30): 7 µg/L; substituted by the MDA value of 14 µg/L in the MPM calculation

15th of November (code 30): 6 µg/L; substituted by the MDA value of 14 µg/L in the MPM calculation

His MPM for 1966 was calculated by SC&A as equal to 89 µg/L.

Worker C had two samples taken on the days around the February incident. Worker C apparently had a much lower intake than the other two, as judged by his excretion results. This worker was monitored only on the 14th and 15th of February. His samples had codes 40 and 49, and his results are described below:

14th of February: 180 µg/L (code 49)

15th of February: 19 µg/L (code 40)

His MPM for 1966 is 100 µg/L.

Worker A and Worker B probably had similar exposures in the February incident. On the first day of sampling (14th of February), both workers presented samples with high concentrations of uranium, about 1,000 µg/L (samples code 40). On the second day of sampling, Worker B had most of the results higher than Worker A. His only sample labeled 40 was similar to the results of Worker A. The follow-up results in February were similar. The 1966 OPOS for those two workers are different though. Worker A had a higher OPOS result because his monitoring frequency was lower than Worker B on the dates following the incident and during the year.

Worker C probably had a lower intake than Worker B in the February incident. His excretion rates in the first day of the incident and in the second day of the incident are lower than the excretion rate of Worker B. The 1966 OPOS result calculated for Worker C was similar to Worker B, because he had a lower frequency of monitoring.

Finding C-3: The annual individual OPOS result has been shown to be influenced by the worker's frequency of monitoring. This is an inherent flaw in the OPOS methodology that cannot be overcome.

C.5 Interpretation of Results that are Less than the Minimum Detectable Level

Table 5-19 of the Fernald site TBD (ORAUT 2004) provides minimum detectable activities/levels for uranium urinalysis by time period; the values in Table 5-19 are replicated below in Table C-5. As seen in the table, an MDA of 14 µg/l is assumed for the entire operational period and part of the residual period at Fernald.

Table C-5: Minimum Detectable Activity/Level for Fernald Uranium Urinalysis Methods

Urinalysis Method	Time Period	Frequency	MDL
Fluorophotometry	1952–1993	Weekly to annual – job specific	14 µg/l*
Chemchek KPA	1993–Sept. 2002	Bimonthly	0.17 µg/l
ICP-MS	Sept 2002–Present	As requested	0.15 µg/l

* Y-12 listed a sensitivity of 1.6 µg/l in 1973 using the fluorometric process for 0.7% U-235. Fernald frequency listed less than 0.003 mg/l in the bioassay data reports. Several blank samples on intercomparison studies also list results as 0.003 mg/l. A value of 0.008 mg/l has also been quoted in the records as the MDL. However, a formal response on January 21, 1993 (Blalock 1993), to a deficiency in the ability to detect 100 mrem CEDE with the existing 0.014 mg/l MDA is accepted as the most reliable representation for historical MDAs for this analytical procedure.

Source: ORAUT 2004, Table 5-19

The urinalysis results in the HIS_20 database are largely "uncensored" in that they are reported even if the numerical quantity of the sample is below the MDA. The number of bioassay samples that were reported below the MDA is shown in Figure C-4. As seen in the figure, a large proportion of the available bioassay samples were reported as values below the assumed MDA. Beginning in the 1970s more than 80% of the available bioassay samples were below the MDA. Figure C-5 further breaks down these below MDA samples into "negative", "zero" and "between zero and the MDA." The number of samples that were reported as negative or zero was generally less than 10% for most periods. Negative results generally weren't observed until the 2nd quarter of 1997. All below-MDA samples appear to have been used "as is" by NIOSH in

calculating their OPOS values and were not truncated or censored at the assumed MDA to yield the "maximum possible mean," (MPM) as espoused in ORAUT-RPRT-0053, *Analysis of Stratified Coworker Datasets*, Rev. 1 (ORAUT 2012b). SC&A noted the same shortcoming in our review of the thorium and Am/Cm/Cf coworker models for SRS (SC&A 2013), as stated in Appendix B of this report. Based on NIOSH's own stated methodology of using the MPM in ORAUT 2012b, it is SC&A's position that it is inappropriate to use numerical quantities that are below the MDA without an adjustment. The issue of MPM in OPOS is discussed in detail in Section 3.2 of the main report.

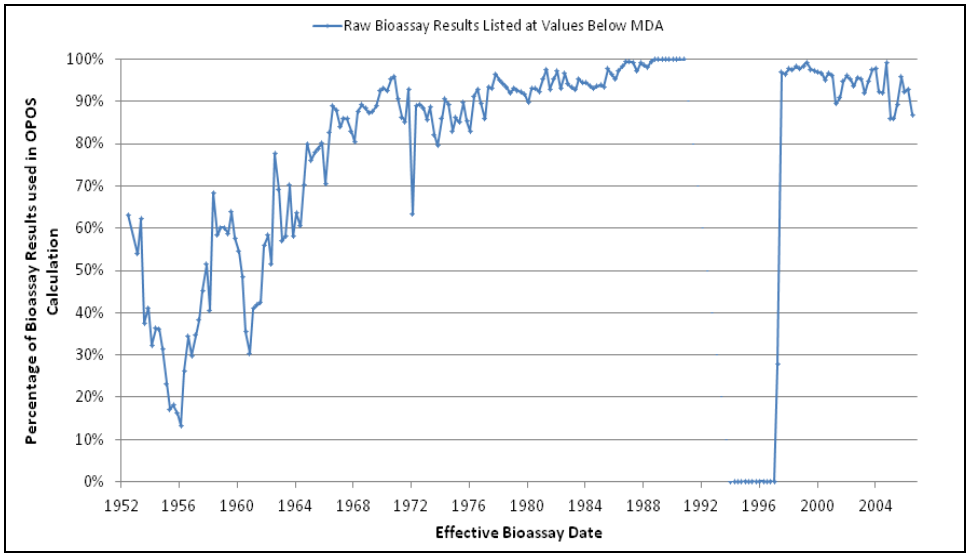

Figure C-4: Percentage of Raw Bioassay Samples Reported Below the Assumed MDA

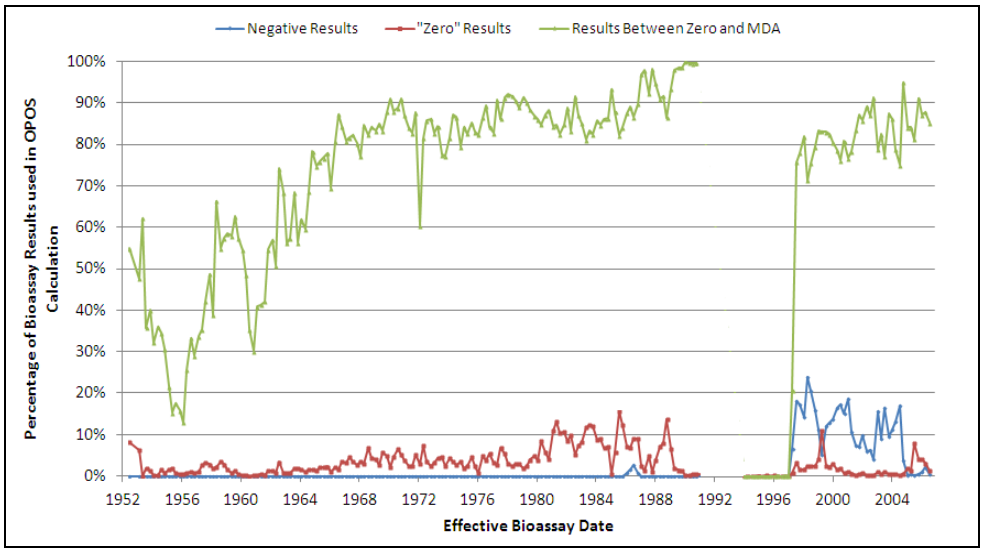

Figure C-5: Less than MDA Bioassay Samples Grouped by Category

Finding C-4: All below-MDA samples appear to have been used "as is" by NIOSH in calculating their OPOS values and were not truncated or censored at the assumed MDA to

yield the "maximum possible mean," (MPM) as espoused in ORAUT-RPRT-0053, *Analysis of Stratified Coworker Datasets*, Rev. 1 (ORAUT 2012b). This same issue was discovered in our reviews of the radionuclide-specific coworker models for SRS.

References

Blalock 1993. *Corrective Action Reports on Fluorometric Fusion Method for Analyzing Uranium in Urine*. Fernald Site – Corrective Action Report, Revision 0. Car Number: 93:0008. ES&H, Radiological Control, Dosimetry. January 21, 1993.

ORAUT 2004. *Technical Basis Document for the Fernald Environmental Management Project (FEMP) – Occupational Internal Dose*. ORAUT-TKBS-0017-5, Rev. 00. Oak Ridge Associated Universities Team, Cincinnati, Ohio. May 28, 2004.

ORAUT 2010. *Internal Dosimetry Coworker Data for the Fernald Environmental Management Project*. ORAUT-OTIB-0078, Rev. 1. Oak Ridge Associated Universities Team, Cincinnati, Ohio. June 3, 2010.

ORAUT 2012a. *Internal Dosimetry Coworker Data for the Fernald Environmental Management Project*. ORAUT-OTIB-0078, Rev. 2. Oak Ridge Associated Universities Team, Cincinnati, Ohio. December 12, 2012.

ORAUT 2012b. *Analysis of Stratified Coworker Datasets*, Rev. 1. ORAUT-RPRT-0053, Oak Ridge Associated Universities Team, Cincinnati, Ohio. July 2012.

SC&A 2013. *SC&A Review of Addendum 3 to the NIOSH Savannah River Site Special Exposure Cohort (SEC-00103) Evaluation Report*. SC&A, Inc., Vienna, Virginia. September 12, 2013.

APPENDIX D: WEIGHTED LEAST SQUARES SOLUTIONS

The linear regression model is expressed concisely in matrix form $\underline{y} = X\underline{ß} + \underline{e}$, where $\underline{ß} = (ß_1, \ldots, ß_k)^T$ is a $(k \times 1)$ vector of regression coefficients, $\underline{e} = (e_1, \ldots, e_n)^T$ is an $(n \times 1)$ vector of independent error terms and, X is an $(n \times k)$ matrix of explanatory variables, and $\underline{y} = (y_1, \ldots, y_n)^T$ is the $(n \times 1)$ vector of responses. Here the symbol $(\ldots)^T$ denotes the matrix transpose. The error terms e_j are independent but are assumed to have unequal variances:

$$Var(e_j) = \frac{\sigma^2}{w_j}$$

where the w_j are weights reflecting the relative precision of the error terms.

The weighted least squares (WLS) solution[4] also is expressed in matrix form:

$$\underline{\hat{\beta}} = [X^T W X]^{-1} [X^T W \underline{y}]$$

Where X is a matrix of explanatory variables, W is a diagonal matrix containing the weights, and $\underline{y}$ is a vector containing the observed values of the response variable corresponding to each row of the X matrix.

For regression through the origin (RTO) with only one explanatory variable, there is only one coefficient $ß$ and the matrix X reduces to a vector of explanatory variables $\underline{x} = (x_1, \ldots, x_n)^T$. This simplifies the WLS solution matrix equation to a simple algebraic form:

$$\hat{\beta} = \left[\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right]^{-1} \left[\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right]$$

$$= \frac{\sum_{j=1}^{n} w_j x_j y_j}{\sum_{j=1}^{n} w_j x_j^2}$$

When the WLS solution is applied to estimate a weighted mean, the matrix x reduces to a vector with a value of 1 in each position $\underline{x} = (1, \ldots, 1)^T$. A reduction from matrix to algebraic form confirms that the weighted mean is the solution to the WLS equations:

$$\hat{\mu} = \left[\begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right]^{-1} \left[\begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right]$$

$$= \frac{\sum_{j=1}^{n} w_j y_j}{\sum_{j=1}^{n} w_j}$$

[4] See, for example, *Linear Regression Analysis*, G.A.F. Seber, John Wiley & Sons, 2003.