Response to SC&A's "Review of ORAUT-RPRT-0071 on External Dose Coworker Methodology"

Response Paper

National Institute for Occupational Safety and Health

June 18, 2025

Nancy Chalmers
Oak Ridge Associated Universities Team

Reviewed by John Cardarelli II
Division of Compensation Analysis and Support

INTRODUCTION

This response paper addresses the 10 observations from SC&A's review of the Oak Ridge Associated Universities (ORAU) Team report ORAUT-RPRT-0071, *External Dose Coworker Methodology* [RPRT-0071; ORAU Team (ORAUT) 2015].

The ORAU Team issued RPRT-0071 in July 2015 to introduce a new method, multiple imputation, to be used to replace the previous practice of substituting half of the limit of detection (LOD) for censored results reported as less than LOD (<LOD). In September 2023, SC&A issued SCA-TR-2023-PR071, *Review of ORAUT-RPRT-0071 on External Dose Coworker Methodology* [SC&A 2023]. Note that RPRT-0071 uses the term "coworker" methodology or modeling. RPRT-0071 was written before the Project adopted the term "co-exposure" methodology or modeling. This response paper uses the term "co-exposure"; quotes from RPRT-0071 might contain the term "coworker."

For context, note that Revision 00 of ORAUT-RPRT-0096, *Multiple Imputation Applied to Bioassay Coworker Models* [ORAUT 2019], was issued in January 2019. Revision 01 was issued in June 2021, and this response refers to it [RPRT-0096; ORAUT 2021]. RPRT-0096 is essentially the internal dose counterpart to RPRT-0071, which focuses on external dose. SC&A issued comments on Revision 00 of RPRT-0096 in 2020 [SC&A 2020]. The SC&A memorandum contained no findings or observations for Revision 00. The memorandum was complimentary of the multiple imputation approach, and the conclusion was, "SC&A concludes that it is a mathematically accurate method for assessing censored bioassay data in the absence of other information (such as the actual raw measurements)" [SC&A 2020, PDF p. 9].

Similarly, the conclusion of the SC&A response to RPRT-0071 states that multiple imputation *"*is a credible approach to increasing data utility and improving co-exposure models that face the problem of underlying datasets with missing and/or censored measurements" [SC&A 2023, PDF p. 14]. SC&A goes on to say that the observations deal with "further exploration of some of the finer points" of multiple imputation [SC&A 2023, PDF p. 14].

SC&A OBSERVATIONS WITH NIOSH RESPONSES

In the sections below, the National Institute for Occupational Safety and Health (NIOSH) presents SC&A's 10 observations with responses after each.

SC&A Observation 1: RPRT-0071 does not include estimates of uncertainty

> *The authors do not capitalize on the benefit of MI* [multiple imputation] *related to estimating uncertainty, a benefit that should be exploited for better understanding of the estimates generated by the MI method. MI can not only help researchers understand the uncertainty involved in making imputations but also help clarify the uncertainty of inferences in downstream methodology, such as co-exposure models and POC* [probability of causation] *calculations* [SC&A 2023, PDF p. 7].

NIOSH Response to SC&A Observation 1

NIOSH agrees that one of the benefits of multiple imputation is the ability to estimate uncertainty. However, as footnote 4 of RPRT-0071 mentions, "the Project does not use the variance in the parameters for anything in the coworker model" [ORAUT 2015, PDF p. 8]. The purpose of RPRT-0071 was to illustrate how multiple imputation compares to substitution of half the limit of detection (LOD/2) in the co-exposure framework, which does not take the uncertainty of the estimates into account.

SC&A Observation 2: RPRT-0071 should expand its exploration of mixture models

> *SC&A would like to see further exploration of issues related to nonpositive measurements, as we believe it relates to all reported measurements, not just the nonpositive ones. A later report, ORAUT-RPRT-0096, revision 01 (ORAUT, 2021; "RPRT-0096"), noted that nonpositive results come from "noise generated when samples containing approximately the same levels of uranium are subtracted from each other" (p. 13). In practice, this type of measurement error is not present just in the nonpositive results; it is there in all the observations. In the same report, Oak Ridge Associated Universities Team (ORAUT) details a possible solution: mixture models (ORAUT, 2021, section 5.0). SC&A believes the development of mixture models is worth further exploration as a fundamental issue in dosage measurement that could potentially be exploited to develop better inferences* [SC&A 2023, PDF p. 7].

NIOSH Response to SC&A Observation 2

NIOSH addresses the "measurement error" portion of this observation in the response to Observation 10. The focus of Observation 2 is the desire to explore mixture models. The purpose of RPRT-0071 was to illustrate how multiple imputation compares to substitution of LOD/2 in the co-exposure framework. The use of mixture models is beyond the scope of RPRT-0071, but ORAUT-RPRT-0110, *Discussion of Variability in Health Physics Data*, explores the use of mixture distributions [ORAUT 2025].

SC&A Observation 3: Determine the appropriate statistical distribution to use for censored readings in each case individually

> *SC&A reiterates a point the authors make in passing: The lognormal distribution highlighted in the report on which to base the MI method is not going to be optimal in all situations. Each situation should be evaluated individually to determine the most appropriate underlying distribution to use for censored readings. It is important for analysts to understand that misspecification of an underlying distribution will undermine the benefits of the MI method* [SC&A 2023, PDF p. 7].

NIOSH Response to SC&A Observation 3

NIOSH concurs. The choice of imputation distribution is made by an ORAU Team statistician and discussed with and reviewed by at least one other ORAU Team statistician and by health physicists. A lognormal distribution is usually appropriate, but the statistician can (and will) consider other distributions if necessary.

SC&A Observation 4: The need to account for relationships between dose and covariates should be considered

> *The primary analysis question is sometimes broader than simply determining an underlying distribution to use in the MI procedures. In fact, there may be situations where accounting for the relationship of dose to other variables is more important than the choice of statistical distribution: If dosage varied by how closely an employee worked to the source, it may be more pressing to use a regression model that makes use of covariate data (e.g., job type) than to empirically fit a sitewide statistical distribution. In such situations, the distribution model might be secondary to the need to account for existing relationships to other variables. One can still assume an underlying lognormal model, for instance, and fit a generalized linear model with covariate data under that assumption for the purposes of MI* [SC&A 2023, PDF p. 7].

NIOSH Response to SC&A Observation 4

The dataset used in RPRT-0071 only contains dates, ambiguous worker identification numbers, and doses. Typically, datasets do not contain information to enable the use of covariates. The statistician doing the analysis can consider covariates, if they are available.

SC&A Observation 5: NIOSH does not provide adequate information on how the RPRT-0071 table 1-1 doses were reconstructed

> *How the doses in table 1-1 of RPRT-0071 were reconstructed seems an important point in assessing the accuracy of an imputation model for these data. SC&A assumes that a raw dataset was available for the example worker. However, further explanation regarding these "reconstructed" doses would be helpful. It would be appropriate for the authors to explain how the doses were reconstructed and the effect of the reconstruction on the multiple imputation model. Of particular interest would be the bias and precision of the reconstruction method and the implications for later inferences of the co-exposure models* [SC&A 2023, PDF p. 8].

NIOSH Response to SC&A Observation 5

The dataset provided to the authors of RPRT-0071 contained an "Actual" values column, where results <LOD had been recalculated based on the number of tracks, blank subtraction, and the

cycle-specific calibration. The raw data were not made available to the authors. The regression on order statistics used for the imputation model in Figure 3-1 does not use the recalculated values that are <LOD, so the imputation model is not affected by how the doses were recalculated.

SC&A Observation 6: RPRT-0071 would benefit from a disclaimer in the discussion of linear imputation

> *Putting the "linear imputation" method in the form of an equation makes it look like a model. In fact, the authors state at the end of RPRT-0071, section 1.0, that because this method "introduces the idea of using a distribution model" (ORAUT, 2015, p. 6) for imputation, they will move on to talking about the statistically based multiple imputation model. This might be an effective visual tool for elucidation, but we worry that someone could read this development of the linear imputation model and think that it is a valid imputation method. Perhaps the authors should add a disclaimer to their development of this model out of caution* [SC&A 2023, PDF p. 9].

NIOSH Response to SC&A Observation 6

NIOSH agrees that the linear imputation method is not an appropriate imputation model. Early in the Project, it was an option for internal co-exposure modeling [ORAUT 2006, PDF p. 8] but has since been replaced with methods described in RPRT-0096 [ORAUT 2021]. It was included in RPRT-0071 for illustrative purposes and as a logical transition from substitution of LOD/2 to an imputation model. A disclaimer could be added in a future revision of RPRT-0071, but the statisticians who perform multiple imputation are aware that the linear imputation method is not appropriate and is no longer used on the Project.

SC&A Observation 7: RPRT-0071 should acknowledge the impact of clustering

> *We would suggest that the authors acknowledge this possible limitation and add a note to their report that the statistician working on a project with clustered data should evaluate the potential impact of clustering on their analysis. If clustering effects are potentially large enough to have a material effect on model fit, the statistician should apply a model fitting method that accounts for the clustering* [SC&A 2023, PDF p. 9].

NIOSH Response to SC&A Observation 7

The purpose of RPRT-0071 was to illustrate how multiple imputation compares to substitution of LOD/2 in the co-exposure framework. Consideration of clustering or stratification is beyond the scope of this report. Clustering and stratification will be considered by a team of statisticians and health physicists on a case-by-case basis.

SC&A Observation 8: RPRT-0071 should provide advice about fitting data that are not lognormal

> *While only based on an example dataset, RPRT-0071, section 3.0, would benefit from more transparency. As a guide for how a statistician should apply the MI method recommended in this report, it would be worthwhile to present a fuller analysis of this particular dataset as a case study. It would also be useful to hear the authors' advice on what to do when lognormality of the data cannot be assumed* [SC&A 2023, PDF p. 11].

NIOSH Response to SC&A Observation 8

The purpose of RPRT-0071 was to illustrate how multiple imputation compares to substitution of LOD/2 in the co-exposure framework. Many exploratory analyses could be done with these data, but that is beyond the scope of this report. If the data are not lognormal, the statistician performing the analysis can use whatever distribution is most appropriate.

SC&A Observation 9: RPRT-0071, section 3.0, should expand its discussion of population subsets

> *Since RPRT-0071 is intended to address procedures in many different situations, it should note that an important potential application is one in which populations of workers differ by level of exposure and those populations may be distinguished by available information, or covariate data. For instance, in DCAS-IG-006, revision 00 (NIOSH, 2020), the National Institute for Occupational Safety and Health (NIOSH) discusses the use of stratification to allow for analysis of highly exposed populations separately from other populations. Such a procedure is potentially a simple and effective way to improve imputations if the covariate data are available to stratify the populations. In some cases, for example, knowledge of job type could be helpful in predicting dosage. In fact, the note about subsetting (ORAUT, 2015, p. 8, third bullet) is an example of how the use of covariate data might be helpful. Instead of subsetting the data by occupation, which would yield a smaller sample size for modeling each occupation, it would perhaps be more effective statistically to use the covariate data related to occupational potential risk in a single (e.g., regression) model that includes all the dosage data to generate the multiple imputations* [SC&A 2023, PDF p. 11].

NIOSH Response to SC&A Observation 9

The purpose of RPRT-0071 was to illustrate how multiple imputation compares to substitution of LOD/2 in the co-exposure framework. NIOSH agrees that there are situations where considering covariates or stratification would be helpful. As mentioned in the response to Observation 4, this dataset did not contain information that would enable consideration of covariate data or stratification. Consideration of covariate data or stratification is beyond the scope of this report.

In general, covariates and stratification will be considered by a team of statisticians and health physicists on a case-by-case basis.

SC&A Observation 10: RPRT-0071 does not acknowledge positive measurement error

> *Given that measurement error is present in all reported dosimeter readings, not just the negative ones, it should be clear that some measurements reported as below the LOD come from doses that are actually above the LOD (negative measurement errors) and that some doses measured as above the LOD come from actual doses below the LOD (positive measurement errors). This means that dose measurements with negative measurement errors are more likely to be imputed than those with positive measurement errors, which is a potentially biased application of imputation* [SC&A 2023, PDF pp. 13–14].

NIOSH Response to SC&A Observation 10

NIOSH disagrees with the interpretation in Observation 10. SC&A says, "The fact that we see nonpositive observations in dosage measurements is due to the presence of measurement error" [SC&A 2023, PDF p. 13]. Measurement error does not cause so many of the measured results to be negative. RPRT-0071 does not address measurement error.

As mentioned in the response to Observation 5, background or blank subtraction was done when the <LOD doses were recalculated. The blank subtraction is causing the negative measured results. That phenomenon necessitated RPRT-0071 footnote 1, which stated that "The true doses are not negative, the measured doses are negative. However, this technicality has no further bearing on this discussion" [ORAUT 2015, PDF p. 6].

SC&A also mentions this quote from RPRT-0096: *"*"The normal component of the mixture can be viewed as the analytical noise generated when samples containing approximately the same levels of uranium are subtracted from each other" [ORAUT 2021, PDF p. 14].

The quote from RPRT-0096 refers to blank subtraction as well. If a bioassay sample that is noise (essentially no activity in it) has a blank subtracted from it, it likely contributes to the normal component of the mixture that is being discussed in the quote. Neither of these reports attributes negative measured results to measurement error.

Measurement error exists anytime a measurement is made. Because of measurement error, measured doses that are <LOD could have true doses above LOD, and measured doses above the LOD could have true doses that are <LOD. This is true in every analysis and is typically not addressed.

Measurement error is much less consequential than Observation 10 makes it seem. A correct understanding of what is causing the nonpositive results (blank subtraction, not measurement error) makes this observation immaterial.

CONCLUSION

SC&A's previous review of RPRT-0096 and this review of RPRT-0071 are complimentary of the multiple imputation method. The 10 observations for RPRT-0071 deal with exploring some of the finer points of multiple imputation. Many of the observations address issues that are beyond the scope of RPRT-0071 but are useful to consider on a case-by-case basis when multiple imputation is used by statisticians.

REFERENCES

ORAUT [2006]. Generating summary statistics for coworker bioassay data. Oak Ridge, TN: Oak Ridge Associated Universities Team. ORAUT-PROC-0095 Rev. 00, June 05. [SRDB Ref ID: 73397]

ORAUT [2015]. External dose coworker methodology. Oak Ridge, TN: Oak Ridge Associated Universities Team. ORAUT-RPRT-0071 Rev. 00, July 02. [SRDB Ref ID: 145135]

ORAUT [2019]. Multiple imputation applied to bioassay coworker models. Oak Ridge, TN: Oak Ridge Associated Universities Team. ORAUT-RPRT-0096 Rev. 00, January 24. [SRDB Ref ID: 175396]

ORAUT [2021]. Multiple imputation applied to bioassay co-exposure models. Oak Ridge, TN: Oak Ridge Associated Universities Team. ORAUT-RPRT-0096 Rev. 01, June 11. [SRDB Ref ID: 187045]

ORAUT [2025]. Discussion of variability in health physics data. Oak Ridge, TN: Oak Ridge Associated Universities Team. ORAUT-RPRT-0110 Rev. 00, May 22. [SRDB Ref ID: 207804]

SC&A [2020]. SC&A memo: review of multiple imputation methods applied to censored bioassay datasets. Memorandum to Savannah River Site and SEC Issues Work Group. Arlington, VA: SC&A. June 3. [SRDB Ref ID: 182225]

SC&A [2023]. Review of ORAUT-RPRT-0071 on external dose coworker methodology. Arlington, VA: SC&A. SCA-TR-2023-PR071 Draft Rev. 0, September 29. [SRDB Ref ID: 207803]