Draft

Advisory Board on Radiation and Worker Health
National Institute for Occupational Safety and Health

Review of ORAUT-RPRT-0071 on External Dose Coworker Methodology

Contract No. 75D30119C04183
Document No. SCA-TR-2023-PR071, Revision 0

Prepared by

Richard Griffiths

SC&A, Inc.
2200 Wilson Blvd., Suite 300
Arlington, VA 22201-3324

September 29, 2023

SC&A, Inc. technical support for the Advisory Board on Radiation and Worker Health's review of NIOSH dose reconstruction program

Document title	Review of ORAUT-RPRT-0071 on External Dose Coworker Methodology
Document number	SCA-TR-2023-PR071
Revision number	0 (Draft)
Supersedes	NA
Effective date	September 29, 2023
Task manager	Kathleen Behling [signature on file]
Project manager	Bob Barton, CHP [signature on file]
Document reviewer(s)	Kathleen Behling [signature on file] Bob Barton, CHP [signature on file]

Record of revisions

Revision number	Effective date	Description of revision
0 (Draft)	9/29/2023	Initial issue

Table of Contents

Abbreviations and Acronyms

CL	censoring level
LOD	limit of detection
MDA	minimum detectable activity
MI	multiple imputation
NIOSH	National Institute for Occupational Safety and Health
ORAUT	Oak Ridge Associated Universities Team
POC	probability of causation

1 Introduction and Background

ORAUT-RPRT-0071, revision 00, "External Dose Coworker Methodology" (ORAUT, 2015; "RPRT-0071"), describes a multiple imputation (MI) procedure that the authors recommend using to replace the current method for imputing censored dosimeter readings. Currently, for a dose reported as censored at less than the limit of detection (LOD), a working estimate of one-half the LOD (LOD/2) is imputed. Using a data-based example, RPRT-0071 examines the benefits of using MI for censored observations and suggests that it can be used in developing complete datasets to support co-exposure models for unmonitored workers.

The MI method, in contrast to the current method, fills in censored measurements with several "replicate" observations. The number of replicate observations can be varied, depending on the application. Once the censored data have been imputed multiple times, an average of the replicate imputations may be taken and used in place of the censored doses. The average imputed doses can then be combined with the uncensored data in further analysis. RPRT-0071 demonstrates the improvement this can yield over the current method using the example data.

As detailed in RPRT-0071, the imputation procedure is composed of two distinct components. The first component is a probability model that describes the distribution of the uncensored dosimeter readings. The second component is the MI method itself. The probability model underpins the MI method by providing a way to replace censored data with observations randomly drawn from the modeled distribution. After the probability model is fit to the uncensored data, the MI method takes multiple random draws from the fitted model for each censored observation.

2 Summary of SC&A's Review

SC&A believes the use of MI, with a probability model serving as the basis for the imputations, for censored dosimeter readings is justifiable and likely a statistical improvement on using LOD/2 to replace censored values. Careful choice and fitting of an underlying probability model, combined with MI, should be effective in reducing the bias of censored doses imputed with the LOD/2 method.

MI is generally regarded as a state-of-the-art method for estimating missing data. This is primarily because MI does a better job of reducing bias than do many traditional imputation methods. Further, MI allows for proper accounting of the uncertainty of estimates generated by the imputation procedure and thus should yield a firmer basis for understanding the statistical error in final probability of causation (POC) estimates.

RPRT-0071 makes use of two distinct lines of thought that need to be understood separately. In addition to the idea of using MI, the use of a probability model to support the MI plays a large role in the presentation. In particular, the use of the lognormal probability model underpins the presentation of the MI method in the report. We view the potential of the MI method positively but note that the application of the lognormal model can be problematic in certain situations and needs to be validated for each individual analysis.

Overall, our view of RPRT-0071 is positive, but we note that there are several broad topics related to MI that could be further explored to improve the imputation of censored dose measurements. These broad topics are encompassed in the following four observations.

Observation 1. RPRT-0071 does not include estimates of uncertainty

The authors do not capitalize on the benefit of MI 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 calculations.

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.

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.

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.

While we agree with a previous SC&A review of MI (SC&A, 2020, p. 6) "that the use of multiple imputation in evaluation of bioassay datasets with censored results is technically appropriate, scientifically defensible, and likely of small practical significance when considering

its effect on resulting POC calculations," in most cases, we also believe that there are several topics that could be further investigated to improve the practical application of MI.

3 SC&A's Comments on RPRT-0071 Discussions and Documentation

3.1 Notes on the introduction

3.1.1 Reconstructed doses

The second paragraph of RPRT-0071, section 1.0, states, "All of the doses that were reported by the site, including those in Table 1-1, were reconstructed to eliminate the censoring. These uncensored doses are given in the Actual column in Table 1-1" (ORAUT, 2015, p. 5). The authors do not explain how the doses were reconstructed. Later in section 1.0, the authors state that "the Actual or Positive doses are seldom if ever known" (p. 5). We are left to wonder if they were reconstructed via a model or if these were some other approximations of the actual recordings.

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.

3.1.2 Negative dose measurements

The authors indicate in footnote 1 on page 5 of RPRT-0071 (ORAUT, 2015) that negative dose measurements come from measurement error. This indicates that a comprehensive statistical model for imputation of dosages should account for measurement error. We feel this is an important topic that would benefit from further exploration. We review this topic more fully in section 4.2 of this report.

3.1.3 The linear imputation model

In RPRT-0071 (ORAUT, 2015, p. 5), the authors state that "These linearly imputed doses are given in the Impute C column in Table 1-1." It is not clear from the text, but what this method entails is the following:

- Take the x-axis of a graph to be the dates of the measurements.
- Take the y-axis of the same graph to be the imputed measurement for each dose.
- Draw a line starting at y = 0 for the first date to y = 0.05 (LOD) for the last date.
- Impute the value of y for the measurement for each date on the x-axis.

So, the imputed values start at 0 for the first date in table 1-1 and end at 0.05 for the last date in table 1-1. A rough equation for this "linear imputation" is as follows:

$$\hat{y} = 0.05 \times t$$

where t represents the percent of the year that has elapsed and $\hat{y}$ is the estimated dose.

We think the authors are trying to help the reader visualize the imputation values given in the last column of table 1-1 by presenting a graphical analogy. Otherwise, the linear imputation method does not have any factual basis: The way it is used in RPRT-0071, section 1.0, assumes that doses increase linearly throughout a year. There is no justification for why this would be the case.

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.

3.2 Notes on RPRT-0071, section 3.0, imputation models and multiple imputation

3.2.1 Clustering

The authors fit a lognormal distribution to a set of 3,736 doses, most of which were reported as censored. These doses came from 732 different workers. On average, there were thus five dose measurements per worker. From a statistical perspective, then, the observations are clustered by worker and potentially have an intra-worker correlation. For instance, doses may be related to job function and individual workers may have had only one job function.

We cannot tell from the description but would guess that the fitting of the lognormal distribution was done assuming independent and identically distributed observations, which is not true of a clustered data set. In practice, this is important because if the intra-cluster correlation is not small there could be a material impact on the model fit. If the intra-cluster correlation is small, violation of the independent and identically distributed assumption will not have a large effect on model fit.

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.

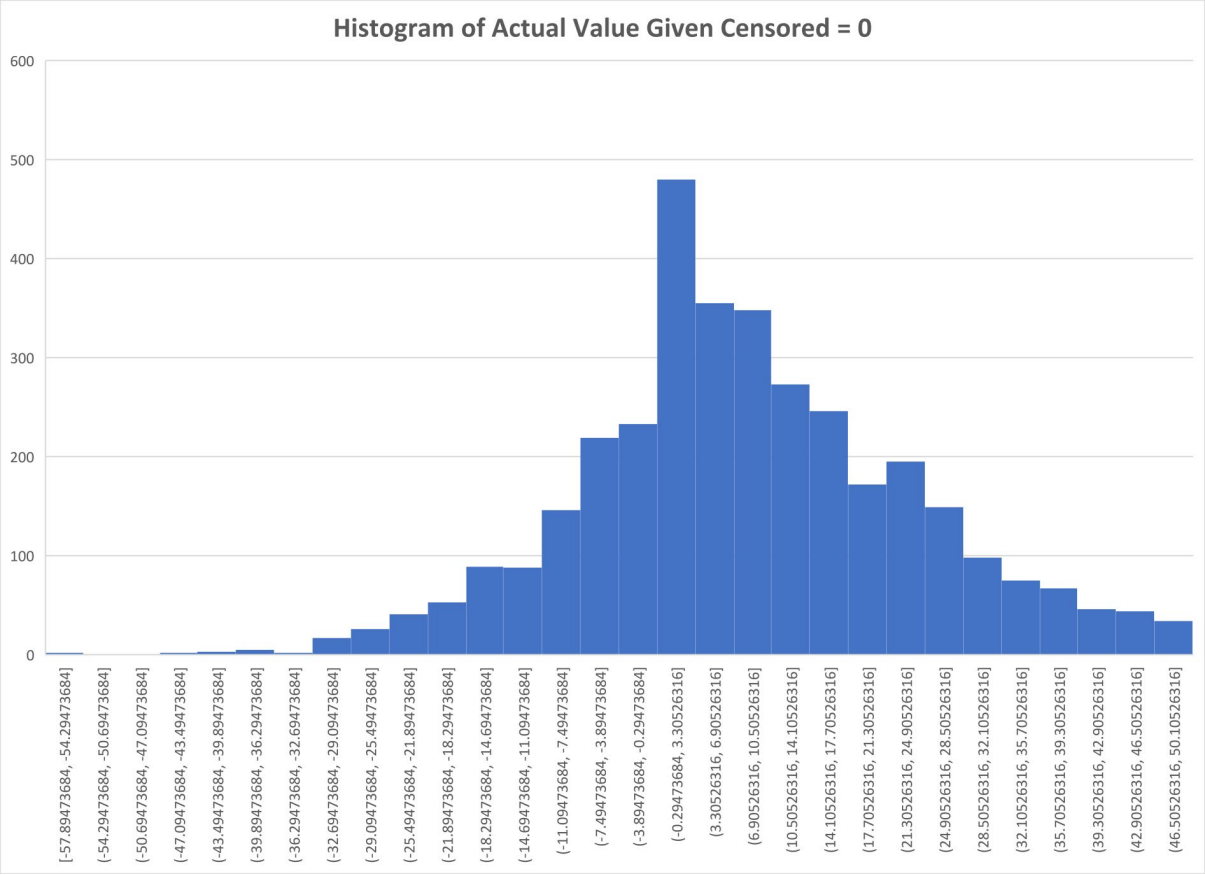
3.2.2 Analysis of data in figure 3-1

The authors of RPRT-0071 (ORAUT, 2015) state that about 94 percent (3,508 out of 3,736) of the observations in figure 3-1 are below the 0.05 rem LOD. This is hard to see in the figure. Since the scale of the figure allows the reader to see the entirety of the data, which range up to

0.25 on the logarithmic scale, the preponderance of observations below the 0.05 level is largely obscured. This is unfortunate because the observations of concern are precisely those observations below the 0.05 level. Per RPRT-0096, "For censored data, the probability distribution chosen for the imputation model should accurately describe the distribution of the data below the CL" (ORAUT, 2021, p. 5.) It is impossible from this figure to tell how well the measurements below the LOD are being described by the lognormal distribution. A better way of graphically depicting the distribution of the smaller observations should be shown. It would be possible, for instance, to include another figure that portrays the same lognormal distribution line but with the y-axis upper limit at 0.05.

Upon examining the data used to construct figure 3-1 of RPRT-0071, we found that the actual measured doses for the 3,508 censored cases produced the histogram in figure 1 of this report. The histogram indicates that these data are distributed more like a normal distribution than the left tail of a lognormal distribution; therefore, it is unlikely that the lognormal model used for MI in this example would result in unbiased imputed values for the censored measurements. It is important to note that this is not a shortcoming of the MI method. It is instead a shortcoming of using the lognormal probability model to support MI in this instance, which highlights the need for an expanded analysis of the underlying probability model used to support the MI method in any given application.

Figure 1. Histogram of actual measurements (censored = 0)



We suspect the observations with values above the LOD are driving the parameter estimates (geometric mean and geometric standard deviation) of the lognormal fit, since more of the variation in the data is in the uncensored points. Because the censored actual measurements appear to follow a distribution other than the lognormal model fit that is driven by the uncensored data, it is possible there is a better model and set of parameter estimates for the observations that will actually be treated by the imputation procedure. Finding that model would provide better imputations in the statistical sense.

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.

3.2.3 Covariate data

On page 8 of RPRT-0071 (ORAUT, 2015), the authors give examples of other methods that can be used to generate MI imputations. The possibility of using covariate data is not mentioned in any of these methods.

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.

3.2.4 Flavors of multiple imputation

MI methods have been applied broadly in many contexts. The one discussed in RPRT-0071 (ORAUT, 2015) is a simplified version of what statisticians would consider the full implementation of MI. Donald Rubin—the originator of the method—advised that the method be implemented within a Bayesian framework (Rubin, 1986). While we think the version of MI considered in this report is appropriate in its context, we caution against conferring all the benefits the MI method has to offer on this particular application.

3.3 Notes on RPRT-0071, section 4.0, coworker modeling

In section 4.0, the authors of RPRT-0071 (ORAUT, 2015) describe the procedure for using the MI methodology to inform a co-exposure model. The authors state that "the statistician performing the analysis will make the judgment as to whether or not a given dataset is large enough to provide usable parameter estimates" (p. 9). The statistician's judgement in this case should not be confined to only an examination of how well the MI model fits but should also include the uncertainty of the parameter estimates. As we noted in observation 1 and further discuss in section 4.1, MI provides an opportunity for the statistician to understand the impact of imputation on co-exposure model and parameter uncertainty. That effect is not explored in RPRT-0071. We feel the report would benefit from a discussion of this topic and could provide guidance to the statistician who uses MI on a dataset that will be used to develop a co-exposure model.

4 Further Research to Improve Co-exposure Models

4.1 How multiple imputation can improve measurement of model uncertainty

MI is widely considered to be one of the best statistical methods currently available for treating missing data problems. MI tends to require fewer assumptions than other imputation techniques. In practice, this results in final estimates less biased than many other methods (Enders, 2010, pp. 1−2). In addition, MI provides an internal mechanism for assessing the uncertainty in the estimates generated via imputation. Measuring the precision of final estimates is difficult with many other imputation methods; the imputations are often treated as deterministic—having been made with certainty. Unfortunately, this is not true, and it leads to an underreporting of the uncertainty in the final estimates.

SC&A thinks the authors miss an opportunity to highlight the importance of understanding the uncertainty in final estimates based on imputation methods, all the more so given that MI is amenable to easy variance estimation calculations. In particular, in RPRT-0071 section 4.0, the authors note how MI can be used to inform a co-exposure model. Using MI in this way has an impact on the uncertainty of the estimated parameters of a co-exposure model and may provide a simpler way of assessing that uncertainty than other methods. This positive should be discussed.

To reinforce this point, we note that the process outlined in section 4.0 could be implemented with $k = 1$ (i.e., single, not multiple, imputation). Doing so would not alter the bias properties of the model: If it is unbiased, it would remain so with $k = 1$. What using $k > 1$ does do, though, is reduce the uncertainty in the final model estimates and provide a method for assessing that level of uncertainty. With $k = 1$, the level of uncertainty is hard to assess. The point here is that the largest benefit of the MI approach with $k > 1$ is not to reduce the bias of the estimates but to reduce (and allow measurement) of uncertainty. The authors should highlight and discuss this benefit more than they do.

Using MI data in co-exposure models allows users to (1) properly account for the extra uncertainty of the model parameters that results from imputation and (2) estimate resultant standard errors of estimates from the models. So, for instance, this could be reflected in confidence intervals calculated from the co-exposure models. Without proper factoring in of the uncertainty from imputations, co-exposure model confidence intervals would tend to understate

upper and lower bounds. Of particular importance, perhaps, is that the upper bound of a 95 percent confidence interval for dosages would be understated.

4.2 Measurement error

The fact that we see nonpositive observations in dosage measurements is due to the presence of measurement error. As explained in RPRT-0096 (ORAUT, 2021, p. 13), the measurement error component comes from "noise generated when samples containing approximately the same levels of uranium are subtracted from each other."

In practice, measurement error is not present just in the nonpositive results; it is there in all the observations. Thus, each dosimeter reading can be expressed as the sum of two components (a true, underlying dose [μ] and a measurement error [m]):

$$y = \mu + m$$

We would really like to model the true values (μ), but this involves more complex methods than fitting a single probability distribution to the observed data. A possible solution is detailed in section 5.0 of RPRT-0096 (ORAUT, 2021): mixture models.

Thinking about the nonpositive y values, then, we note that the true value (μ) must be greater than or equal to zero. Therefore, any nonpositive observation (y) results from a negative measurement error (m). So, the nonpositive results are not random—they are associated with negative measurement errors. A bias in the model will result if we force the model to fit only positive results (i.e., if we use a model based solely on a non-negative distribution) without properly accounting for the measurement error component. The mixture model of RPRT-0096 (ORAUT, 2021, section 5.0) attempts to overcome this bias. It would seem a similar model could be applied fruitfully in the situation described in RPRT-0071 (ORAUT, 2015) to overcome the bias.

To be clear, the authors of RPRT-0071 (ORAUT, 2015) leave that possibility open in section 3.0 when describing implementing an imputation model in practice (the first bullet point on p. 8). However, the prescription of section 4.0 implies the MI procedure for a co-exposure model should be based on a lognormal model. The section 4.0 prescription should include the possibility of mixture models.

4.2.1 Negative doses and doses near the LOD

As part of the discussion of measurement error, we also note that there is a contradiction in imputing non-negative values for negative reported doses and not acknowledging the measurement error of positive reported doses near the LOD.

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.

While the bias may be small, it underscores the idea that measurement error is present in all dose readings; co-exposure models (and other downstream analyses) that treat dose readings as free of measurement error potentially underrepresent the amount of uncertainty in POC calculations.

5 Conclusion

The presentation of MI as a methodology for use in filling in analytically useful data for censored observations in RPRT-0071 (ORAUT, 2015) 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. MI is generally regarded as a state-of-the-art method (Enders, 2010), primarily because it does a better job of reducing bias and allows for proper accounting of the uncertainty of estimates generated by the imputation procedure than do many traditional imputation methods.

While we agree with a previous SC&A review of MI (SC&A, 2020, p. 6) "that the use of multiple imputation in evaluation of bioassay datasets with censored results is technically appropriate, scientifically defensible, and likely of small practical significance when considering its effect on resulting POC calculations," in many cases, we would like to see further exploration of some of the finer points of MI for their potential to benefit claimants through the improvement of co-exposure models, if MI is to be pursued further. These points are summarized in the following 10 observations:

- Observation 1. RPRT-0071 does not include estimates of uncertainty
- Observation 2. RPRT-0071 should expand its exploration of mixture models
- Observation 3. Determine the appropriate statistical distribution to use for censored readings in each case individually
- Observation 4. The need to account for relationships between dose and covariates should be considered
- Observation 5. NIOSH does not provide adequate information on how the RPRT-0071 table 1-1 doses were reconstructed
- Observation 6. RPRT-0071 would benefit from a disclaimer in the discussion of linear imputation
- Observation 7. RPRT-0071 should acknowledge the impact of clustering
- Observation 8. RPRT-0071 should provide advice about fitting data that are not lognormal
- Observation 9. RPRT-0071, section 3.0, should expand its discussion of population subsets
- Observation 10. RPRT-0071 does not acknowledge positive measurement error

6 References

National Institute for Occupational Safety and Health. (2020). *Criteria for the evaluation and use of co-exposure datasets* (DCAS-IG-006, rev. 00). https://www.cdc.gov/niosh/ocas/pdfs/dr/dc-ig-006-r0-508.pdf

Enders, C. K. (2010). *Applied missing data analysis*. The Guilford Press.

Oak Ridge Associated Universities Team. (2015). *External dose coworker methodology* (ORAUT-RPRT-0071, rev. 00).

Oak Ridge Associated Universities Team. (2021). *Multiple imputation applied to bioassay co-exposure models* (ORAUT-RPRT-0096, rev. 01). https://www.cdc.gov/niosh/ocas/pdfs/orau/oraurpts/or-rprt-96-r1-508.pdf

Rubin, D. B. (1986). Basic ideas of multiple imputation for nonresponse. *Survey Methodology*, *12*(1), 37−47. https://www150.statcan.gc.ca/n1/en/pub/12-001-x/1986001/article/14439-eng.pdf?st=Rm-mgRKJ

SC&A, Inc. (2020, June 3). *Review of multiple imputation methods applied to censored bioassay datasets* [Memorandum]. https://ftp.cdc.gov/pub/FOIAREQ/182225-508.pdf