A Comparative Analysis and Decision Framework for Handling Non-Detects in TSCA Risk Evaluations

Bingbing Owen,[1] Benjamin Roberts,[2] A. Michael Ierardi,[1] Andy Maier [1] [1]Integral Consulting Inc.; [2]Benchmark Risk Group

Introduction

The Challenge of Non-Detects

- **Non-detects (NDs)** occur when measured concentrations fall below an instrument's detection limit, complicating exposure assessment in occupational data sets
- TSCA risk evaluations face additional challenges due to high frequency of NDs and complex data structures
- Multiple statistical approaches exist—each with unique strengths and limitations depending on percent censoring, sample size, and data distribution
- No universally accepted guideline exists, resulting in inconsistent practices among industrial hygiene professionals

Study Objectives

- **Catalog existing guidance** on handling NDs in occupational data sets through scoping literature review
- **Evaluate the performance** of identified statistical methods using randomly generated data sets under varying levels of censoring
- **Develop a decision framework** to guide selection of the most appropriate ND-handling method based on data set characteristics

Methods

Scoping Literature Review

- Searched PubMed/NCBI, EPA guidelines, AIHA resources (e.g., Hewett, 2015), and gray literature
- Focused on peer-reviewed articles and agency documents on ND handling

Data Simulation and Comparative Analysis

- Generated synthetic data sets from 250 random values following a log-normal distribution
- Varied censoring levels (low: <15%; moderate: <50%; high: <80%; severe: <95%)
- Applied five censoring tiers to values below the LOD by randomly sampling an LOD from a uniform range defined as (1–5x, 5–10x, 10–15x, 15–20x, and 20–25x) the original detected value (i.e., the simulated true value)
- Applied several ND-handling methods to each scenario
- Calculated the mean and 95% CI for each dataset using the EnvStats and NADA2 packages in R
- Calculated difference in mean and 95% CI estimates relative to true parameters as well as percent difference (Figure 1)

Framework Development

- Constructed a decision framework with key factor considerations
- Mapped data set characteristics to recommended methods

Results

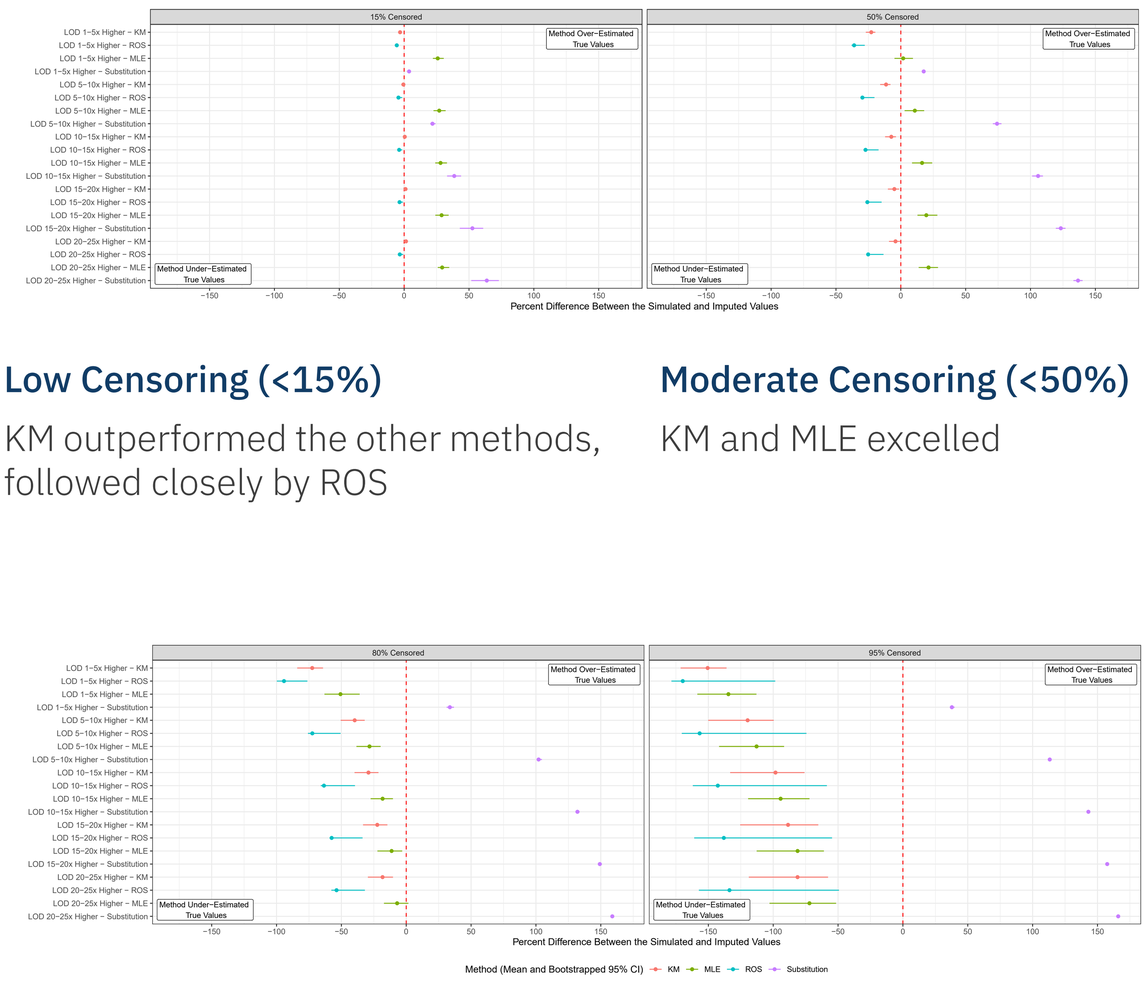
The most widely used statistical methods for NDs are:

- **x=y** Simple substitution
- Maximum Likelihood Estimate (MLE)
- Regression on Order Statistics (ROS) and other regression-based techniques
- Kaplan-Meier (KM)
- KM hybrid
- **β = x** β-substitution
- Multiple imputation (MI)
- Bayesian

Declaration of Interests – All authors are employed by Integral Consulting Inc. or Benchmark Risk Group. No entity outside of Integral or Benchmark funded, reviewed, or provided input on the development of this poster. References available upon request.

Results

Figure 1. Performance Analysis



Low Censoring (<15%)

KM outperformed the other methods, followed closely by ROS

Moderate Censoring (<50%)

KM and MLE excelled

High Censoring (<80%)

MLE seems to perform slightly better than the other methods, with the performance improving as the distance between the LOD and the "true" value increases

Severe Censoring (<95%)

No single method reliably estimates parameters

Framework for Handling NDs in TSCA Risk Evaluations

Based on our review and exploration of the randomly generated data set, we developed a decision framework to guide practitioners in selecting the most appropriate method for handling NDs in a given data set.

Key Data Set Factors

1. Number of Detection Limits
2. Summary Statistics of Interest
3. Sample Size (N)
4. Proportion of NDs
5. Data Distribution
6. Skewness (i.e., GSD)
7. Presence of Outliers

Method Selection Framework

- Low Censoring: KM, ROS
- Moderate Censoring: KM, MLE
- High Censoring: MLE

Notes:

- Generally, simple substitution overestimates the true value in all cases
- No currently recommended methods for severe censoring

Discussion

- Our scoping review confirmed the heterogeneity of existing guidance
- Simulation results quantified method-specific differences
- **Limitations:** Simulated data do not capture all real-world complexities (e.g., mixed distributions, real sampling errors, multiple LODs, varying sample size)
- **Future Directions:** Extend validation to field data sets; refine Bayesian models for severe censoring; develop user-friendly software implementations

Conclusion

- Selecting an ND-handling method requires balancing censoring level, sample size, and statistical goals
- Adoption of this framework can enhance accuracy of exposure estimates, improve risk characterization, and support regulatory compliance
- Establishing consensus guidelines will elevate the rigor and consistency of industrial hygiene practice across diverse occupational settings

Methods used to handle NDs in occupational exposure data sets as part of TSCA risk evaluations are highly sensitive to data set characteristics. Comparison across multiple methods is recommended.

Bingbing Owen, Ph.D., CIH
856.399.5762
bowen@integral-corp.com