Data Completeness in Co-Exposure Models

Brant Ulsh, Ph.D., CHP
Research Health Scientist
Division of Compensation Analysis and Support
NIOSH
February 11, 2026

A Discussion of Completeness in Co-Exposure Models

White Paper

National Institute for Occupational Safety and Health

January 26, 2023

Thomas LaBone
Oak Ridge Associated Universities Team

Reviewed by Tim Taulbee, John Cardarelli, and Lara Hughes
Division of Compensation Analysis and Support

Outline

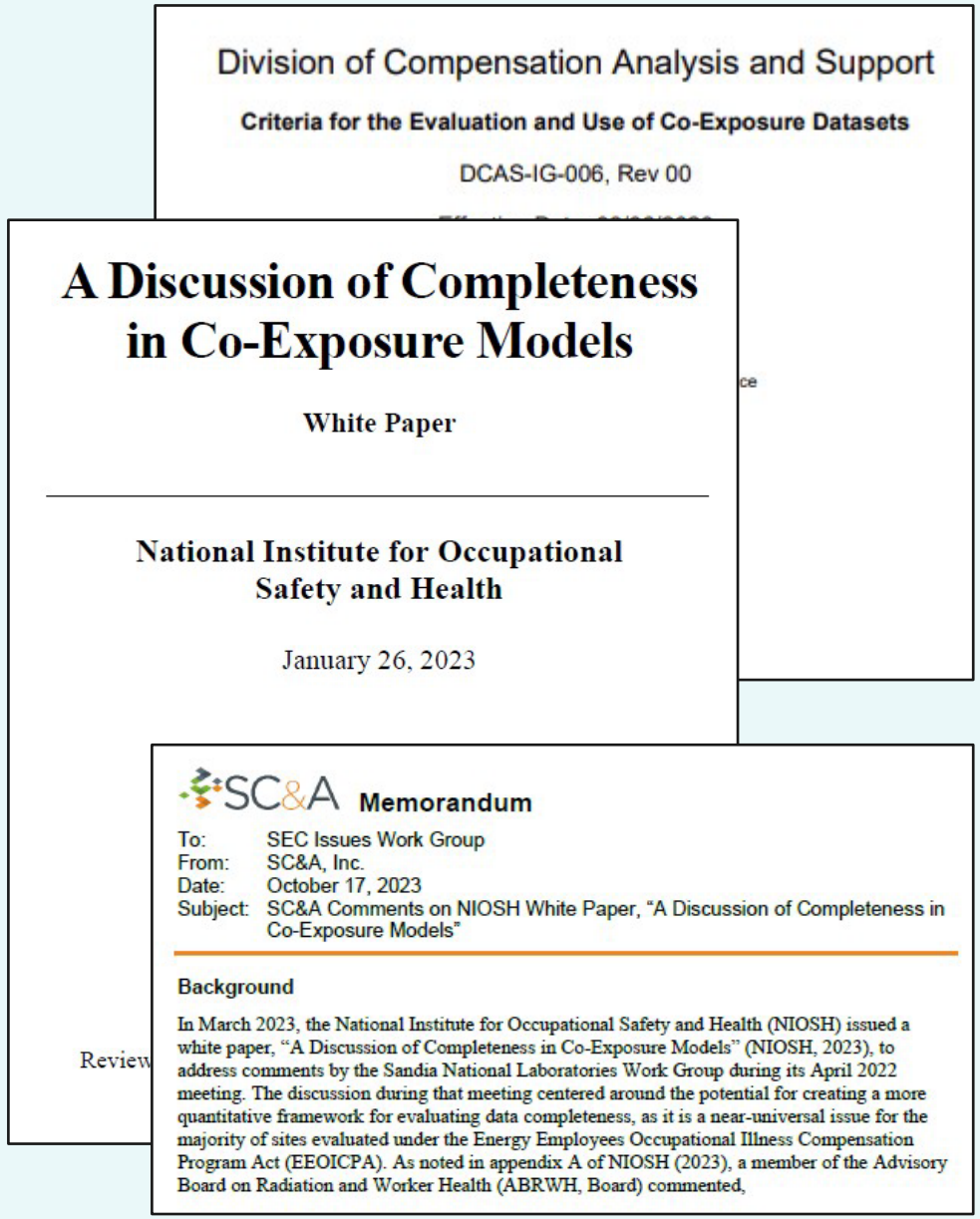
- Brief history – how did we get here?
- Co-exposure models
- Data completeness
- Effects of incomplete data sets
- Starting assumptions
- Regulatory compliance
- Stratification
- Conclusions

History

How did we get here?

Milestones

- DCAS-IG-006, Section 2.2
- 2022-04-11 Sandia National Laboratories Working Group meeting: requested NIOSH prepare report on data completeness
- 2023-01-26 NIOSH/ORAUT issued white paper, “*A Discussion of Completeness in Co-Exposure Models*”
- 2023-10-17 SC&A issued memo, “*SC&A Comments on NIOSH White Paper, “A Discussion of Completeness in Co-Exposure Models*”



Division of Compensation Analysis and Support

Criteria for the Evaluation and Use of Co-Exposure Datasets

DCAS-IG-006, Rev 00

A Discussion of Completeness in Co-Exposure Models

White Paper

National Institute for Occupational Safety and Health

January 26, 2023

SC&A **Memorandum**

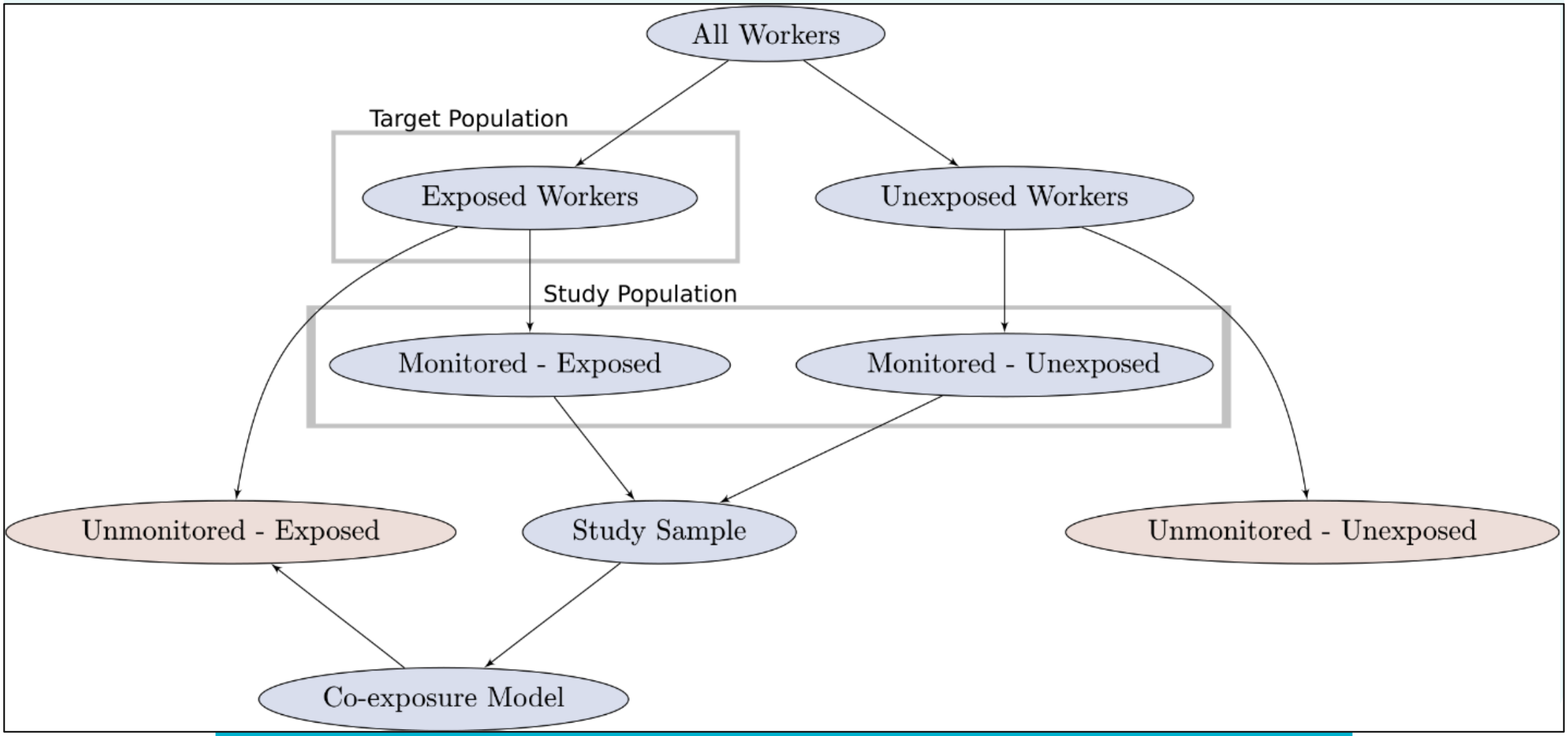
To: SEC Issues Work Group
From: SC&A, Inc.
Date: October 17, 2023
Subject: SC&A Comments on NIOSH White Paper, “A Discussion of Completeness in Co-Exposure Models”

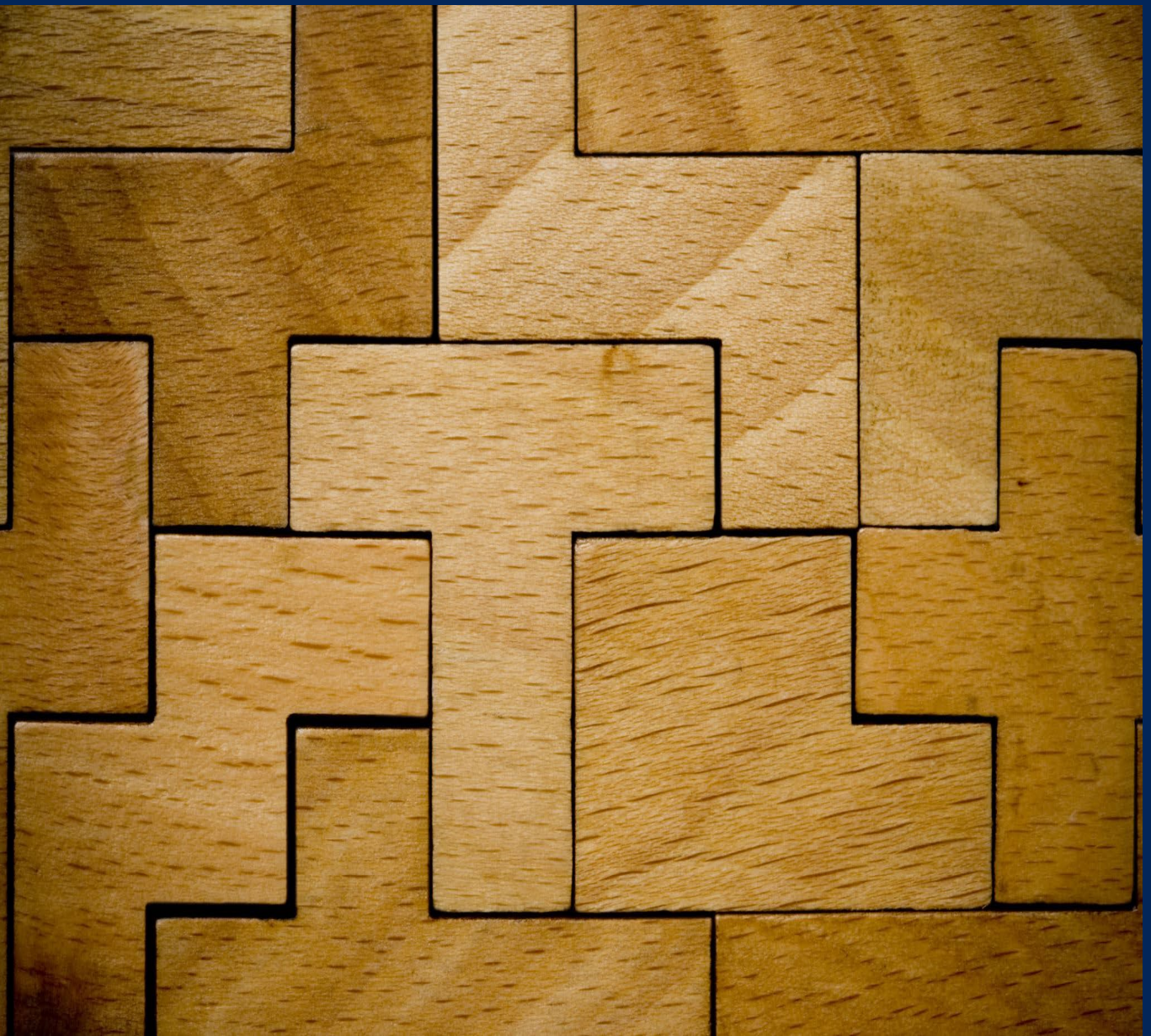
Background

In March 2023, the National Institute for Occupational Safety and Health (NIOSH) issued a white paper, “A Discussion of Completeness in Co-Exposure Models” (NIOSH, 2023), to address comments by the Sandia National Laboratories Work Group during its April 2022 meeting. The discussion during that meeting centered around the potential for creating a more quantitative framework for evaluating data completeness, as it is a near-universal issue for the majority of sites evaluated under the Energy Employees Occupational Illness Compensation Program Act (EEOICPA). As noted in appendix A of NIOSH (2023), a member of the Advisory Board on Radiation and Worker Health (ABRWH, Board) commented,

Co-Exposure Models

Constructing Co-Exposure Models

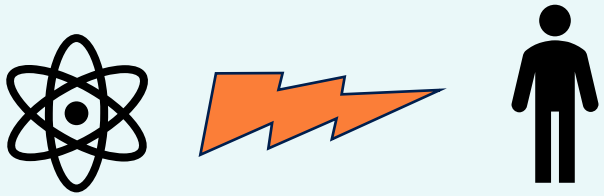




Data Completeness

Data Completeness Basics

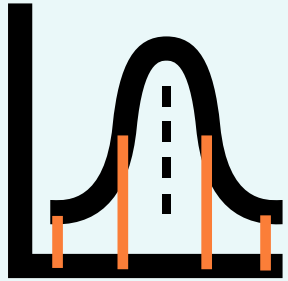
- Co-exposure data sets are **ALWAYS** incomplete – by definition
- Is there enough data from the most highly exposed workers to create a (1) representative, **OR** (2) bounding model?
  - Model doesn't have to be representative if it is bounding.
  - Model doesn't have to be bounding if it is representative.
- The NIOSH/ORAUT white paper introduced the term "missingness" (i.e. opposite of completeness).
- Data sets could be incomplete either because:
  - Workers were monitored, but records are unavailable (i.e. "missing").
  - Workers who may have had an exposure potential were not monitored.



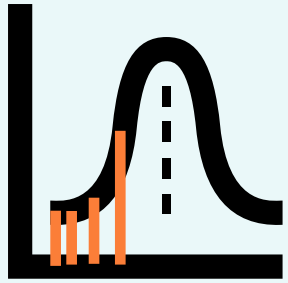
Effects of Incomplete Data Sets

Incompleteness

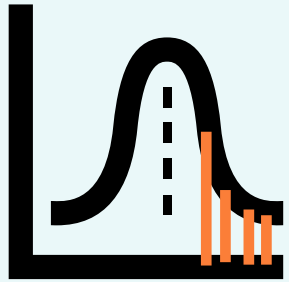
- Is incompleteness a problem?
  - If records are not included at random, model is representative.
    - NOT a problem!

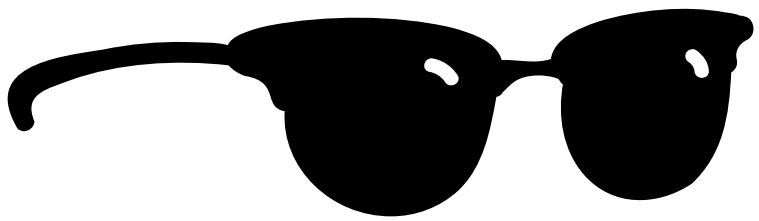


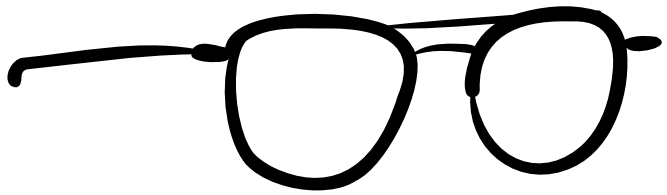
  - If records from low exposed workers are not included, model is biased high.
    - NOT a problem! *(provided model is sufficiently accurate)*



  - If records from high exposed workers are not included, model is biased low.
    - PROBLEM!







Starting Assumptions

Assume Data are Complete?

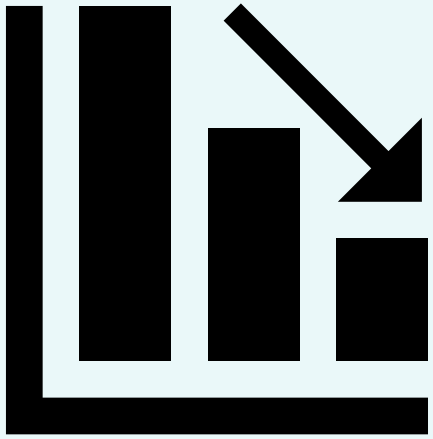
- **NIOSH** (2023): "After vetting the programs and starting under the presumption that the data are complete, a limited number of internal and external checks of the datasets can be performed to look for signs of significant missingness or monitoring missingness".
- **SC&A** (2023): "SC&A does not agree that after vetting appropriate documentation and interviews of the monitoring program that the available dataset should "start under the presumption that the data are complete." SC&A notes that neither should it be assumed that the dataset is incomplete, SC&A believes the evaluation of a given dataset using "internal and external checks" should always start with the neutral assumption that it is not known whether the data are complete or incomplete until such checks are performed".
- Why do NIOSH and SC&A have different starting assumptions?

Basic Scientific Method – Hypothesis Testing

- **Hypothesis testing** is a foundational technique in the Scientific Method.
  - It is not possible to “prove” a hypothesis.
  - The best we can achieve is to rigorously test a hypothesis and fail to disprove it.
    - This provides some confidence in the hypothesis, but it is never proved!
      1. It can always be argued that a more sensitive test would allow us to disprove a hypothesis, especially if that test doesn’t currently exist or is impractical.
    - A single legitimate, validated counter-observation is enough to disprove a hypothesis.
- Defining hypotheses appropriately is critical.
  - Hypotheses MUST be falsifiable, otherwise testing them tells us nothing.
  - **Null Hypothesis**: This is the hypothesis we are trying to **disprove**.
  - **Alternative Hypothesis**: This is the hypothesis of interest.
    - The burden of proof is always on the Alternative Hypothesis to provide an observation that disproves the Null Hypothesis.

Hypothesis Testing – General Example

- General example: A test of a new drug to treat diabetes
  - Null hypothesis: The drug has no (or null) effect on blood sugar.
  - Alternative hypothesis: The drug statistically significantly lowers blood sugar compared to controls who didn't receive the drug.



Hypothesis Testing – Co-Exposure Model Example

- Co-exposure example: A test of whether or not plutonium exposure data from Rocky Flats is complete enough to build a model to use for unmonitored workers.
  - Null hypothesis: The data set is complete.
  - Alternative hypothesis 1: The data set is missing randomly distributed plutonium bioassay data.
  - Alternative hypothesis 2: The data set is missing plutonium bioassay data from workers with low potential exposure.
  - Alternative hypothesis 3: The data set is missing plutonium bioassay data from workers with high potential exposure.
- Why not set the null hypothesis as: The plutonium exposure data from Rocky Flats may or may not be complete?
  - Can’t test it!
- Why not set the null hypothesis as: The plutonium exposure data from Rocky Flats is incomplete?
  - This hypothesis is unfalsifiable. Can always speculate that there might be another box somewhere – especially if no evidence is required to make the claim!
  - The burden of proof MUST fall on the alternative hypothesis!

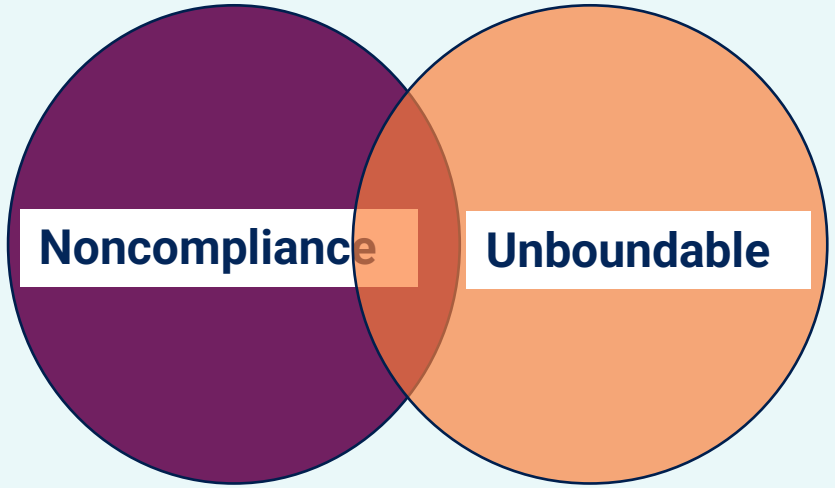
Shifting the Burden of Proof

- Shifting the burden of proof away from the party making a claim onto another party to disprove is also known as “proving the negative”.
  - This is a formally recognized logical fallacy: https://yourlogicalfallacyis.com/burden-of-proof
  - Arguments based on this strategy are invalid because they employ flawed logic.
- Example: we have evaluated the radiation protection program at site XYZ, and it was designed to monitor the highest exposed workers. There were no relevant regulatory violations. We have evaluated the monitoring data and found no evidence that workers with the highest exposure potentials were unmonitored. Therefore, we can bound the dose.
  - You can’t prove that there wasn’t some hypothetical black operation where workers could have received unmonitored exposure, so grant a SEC class.
  - You can’t prove there isn’t some undiscovered cache out there somewhere that might contain monitoring records for workers with even higher exposures.

Regulatory Compliance

Regulatory Compliance and Completeness

- Regulatory compliance is not necessarily related to the ability to construct co-exposure models.
  - Compliance with the regulations does not by itself prove that one can construct a co-exposure model.
  - Noncompliance with the regulations by itself does not prove that one cannot construct a co-exposure model.
  - Regulatory compliance cannot be used by itself to judge the monitoring completeness of a dataset.
  - Regulatory compliance can add to the weight of the evidence, with further evaluation for relevance.



Stratification

Challenges of Stratification

- **If the radiation protection program at a facility was mature and functional**, stratification is not needed. A bounding unstratified co-exposure model is likely to:
  - Assign higher doses to the low-dose workers than would be assigned from a stratified co-exposure model (thus being claimant favorable), and
  - Assign lower doses to the high-dose workers than they would be assigned from a stratified co-exposure model—but those workers are most likely to be monitored and will not need the co-exposure model.
- Stratification when there are in fact no significant differences in exposure potential will produce less accurate co-exposure models because of the reduction in the sample size of each stratum.
- Workers must be accurately assigned to the appropriate strata in the development of the co-exposure model and in the application of the co-exposure model to a specific worker.
- Stratified co-exposure models will result in some workers being assigned more dose and some workers being assigned less dose when compared to the doses assigned from an unstratified co-exposure model.
- Stratification cannot be used to deal with target populations where a particular group of workers is assumed to be the most highly exposed and whose data are not included.

Conclusions

Conclusion

- A co-exposure model does not require all the data, just a significant portion of the data from the most highly exposed workers.
- If a radiation protection program is working properly, workers with high exposure potential are more likely to be monitored, whereas missing data can exhibit practically any pattern.
- Internal and external checks of the datasets can be performed to look for signs of significantly incomplete data or incomplete monitoring.
- It is not feasible to establish universally applicable, technically based quantitative limits for a dataset being complete enough.
  - This is primarily a qualitative decision based on the preponderance of evidence.
- The null hypothesis must be that the co-exposure data set is complete, and we rigorously test to disprove it.
- Regulatory compliance with a monitoring program or lack thereof cannot be used by itself to decide if a dataset is complete enough.
- Stratification of datasets cannot be used to correct for data missingness or monitoring missingness.

Contact Info:

Brant Ulsh, Ph.D., CHP

Research Health Scientist

Division of Compensation Analysis and Support

National Institute for Occupational Safety and Health

Email: bau6@cdc.gov

Thank you.

For more information, contact CDC
1-800-CDC-INFO (232-4636)
TTY: 1-888-232-6348 https://www.cdc.gov/ | https://www.cdc.gov/niosh
Follow us on social **@CDCgov**

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the U. S. Centers for Disease Control and Prevention or the National Institute for Occupational Safety and Health.