Data interpretation and professional judgment in occupational exposure assessment

Gurumurthy Ramachandran, Ph.D., CIH
Division of Environmental Health Sciences, School of Public Health, University of Minnesota
Exposure reconstruction for epidemiological studies

Routine industrial hygiene exposure management
Retrospective Exposure Assessment

A Bayesian Approach to Retrospective Exposure Assessment

Gurumurthy Ramachandran and James H. Vincent

Division of Environmental and Occupational Health, University of Minnesota, Minneapolis, Minnesota
Reconstruct **past** exposures of individuals or groups as a function of time.

<table>
<thead>
<tr>
<th>Species 1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PLANT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JOB 1</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>JOB 2</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>ETC.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**WORKER A**

WORK HISTORY

HEALTH HISTORY

STATISTICAL ANALYSIS

DOSE-RESPONSE RELATIONSHIP

NEW STANDARDS
Uncertainties and Biases

- Sparse historical measurements
- Changes in the environment
  - Changes in workplace practices.
  - Changes in industrial processes.
  - Changes in ventilation patterns.
- Changes in what we measure
  - Changes in measurement criteria.
  - Changes in instruments and analysis methods.
Need additional inputs:

• Historical information about workplace conditions.

• Expert judgments from professionals and researchers with relevant experience/insights.

• Use above information in exposure models.
BAYESIAN RETROSPECTIVE EXPOSURE ASSESSMENT

Incomplete, raw exposure measurements using different methods.

Plant working conditions over its operating history. Industrial processes, production rates, work practices etc.

Normalize all data to personal, inhalable reference.

Experts’ judgment of model parameters.

Models for dust generation, ventilation, and worker activity patterns.

Estimation of personal, subjective biases.

Subjective prior probability distributions of exposure.

Bayesian Update

Estimate variances in exposures for worker groups.

Exposure history and probability distributions of exposure at each point in time, for a given task group.

FIGURE 1
Bayesian methodology for retrospective exposure assessment.
Retrospective Exposure Assessment using Bayesian Methods

GURUMURTHY RAMACHANDRAN*

Division of Environmental and Occupational Health, School of Public Health, University of Minnesota, Mayo Mail Code 807, 420 Delaware Street SE, Minneapolis, MN 55455, USA
BAYESIAN METHODOLOGY

**Historical information** → **Expert inputs** → **Exposure Model Output**

- **Prior (P₀(e))**
- **Posterior (P_{post}(e/M))**

\[
P_{post}(e/M) = \frac{P₀(e)P_L(M/e)}{P(M)}
\]
## Available Measurements

<table>
<thead>
<tr>
<th>Period</th>
<th>Konimeter (ppcc)</th>
<th>HiVol (mg dust/m³)</th>
<th>Personal (mg Ni/M³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1956-1963</td>
<td>959 (24) (765-1153)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1964-1966</td>
<td>561 (15) (458-664)</td>
<td>17 (1)</td>
<td>--</td>
</tr>
<tr>
<td>1967-1971</td>
<td>623 (27) (530-716)</td>
<td>16.5 (2) (0-98)</td>
<td>--</td>
</tr>
<tr>
<td>1976-1979</td>
<td>--</td>
<td>1.31 (17) (0-79-1.83)</td>
<td>4.35 (11) (0-10.4)</td>
</tr>
</tbody>
</table>
Assessing Uncertainty in Measurements

Analytical Variability:
Random measurement errors.
Conversion factors obtained from side-by-side experiments.

Environmental Variability
Spatial and temporal variability
Between- and within-worker variability

Systematic Biases
Concentration history using converted measurements

- Concentration history using converted measurements
- Inhalable Concentration (mg/m³)
- Year
Estimates of exposure with such large error bars are not useful for developing quantitative dose-response relationships for epidemiology.

Need to supplement the sparse data - historical information about workplace conditions, expert judgments, and exposure models.
DIRECT ELICITATION OF EXPERT JUDGMENT

• Ask an expert to directly give us the subjective probability distribution of exposure, i.e., the prior.

• This subjective assessment is a complex decision, resting on a hierarchy of assumptions.

• A direct elicitation ‘hides’ the rationale behind the decision-making process.
HENCE MODELING

It is better to **disaggregate the judgment** by modeling each individual aspect of the problem.

- It produces judgments that are closer to the truth (Morgan and Henrion, 1990).
**Exposure Model**

<table>
<thead>
<tr>
<th>$C_{equil}$</th>
<th>$\frac{G}{Q_{vent}}$</th>
<th>$\frac{P \times E}{Q_{vent}}$</th>
</tr>
</thead>
</table>

| $Q_{vent}$ = Ventilation flowrate (m$^3$/minute) |
| $G$ = Aerosol mass generation rate (mg/min) |
| $P$ = Bessemer matte production rate (tons/minute) |
| $E$ = Emission factor (mg/ton) |
Information Provided to Expert

• Nickel smelting operation description
• Process flow sheets and layout maps
• Ventilation surveys.
• Emission factors.
• Production levels of Bessemer matte (annual)
Monte Carlo Simulations

Probabilities:
- Production Rate (in thousands of tons/year):
  - 314
  - 322
  - 330
  - 339
  - 347
- Flow Rate (in thousands of m³/min):
  - 100
  - 125
  - 150
  - 175
  - 200
- Emission Factor (kg/ton):
  - 0.10
  - 0.15
  - 0.20
  - 0.25
  - 0.30
- Concentration mg/m³:
  - 1.00
  - 2.50
  - 4.00
  - 5.50
  - 7.00

100,000 Trials
Comparison of Expert Opinions

Inhalable Nickel Concentration (mg/m³)

Year


Expert 1

Expert 2
Conventional exposure history vs. Bayesian posteriors

Inhalable Nickel Concentration (mg/m³)

Year


Expert Two
Expert One
Measurements
Bayesian Decision Analysis
IH Effectiveness Goal:

Ensure that no worker has unacceptable exposures
Question:

Most common number of air samples used to make a judgment about exposure?

A. >10
B. 6 to 10
C. 3 to 5
D. 1 or 2
E. 0
Example of Medium Sized Manufacturing Facility

<table>
<thead>
<tr>
<th></th>
<th>Number of Employees</th>
<th>Exposure tasks</th>
<th>Chemicals x tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>200</td>
<td>20</td>
<td>300-400</td>
</tr>
<tr>
<td>Maintenance</td>
<td>40</td>
<td>40</td>
<td>600-800</td>
</tr>
<tr>
<td>Engineering</td>
<td>60</td>
<td>25</td>
<td>350-500</td>
</tr>
<tr>
<td>Admin.</td>
<td>10</td>
<td>2</td>
<td>30-40</td>
</tr>
<tr>
<td>Other</td>
<td>15</td>
<td>8</td>
<td>120-160</td>
</tr>
</tbody>
</table>

Assumption: 15-20 chemicals per task
Professional Judgments: Why Important?

- Leverage Data and Information
- Integrate Wide Range of Inputs
- Help Deal With Uncertainty

Consequences if Wrong:
- Inconsistent Level of Protection
- Wasted Resources
<table>
<thead>
<tr>
<th>Hazardous Materials Management And Procurement</th>
<th>Personal Protective Equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposure Monitoring</td>
<td>Hearing Conservation</td>
</tr>
<tr>
<td>Medical Surveillance</td>
<td>Engineering Controls</td>
</tr>
<tr>
<td>Education And Training</td>
<td>Work Practice Controls</td>
</tr>
<tr>
<td>Training Hazard Communication</td>
<td>Administrative Controls</td>
</tr>
<tr>
<td>Compliance Epidemiology</td>
<td>Environmental</td>
</tr>
<tr>
<td>Exposure Assessment</td>
<td></td>
</tr>
</tbody>
</table>
## Class Survey

Xylene: TLV = 100 ppm

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Data (ppm)</th>
<th>Interpretation - Acceptable?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21, 68</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>21, 109, 38, 41, 48</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>12, 16, 21, 24</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>8, 70, 5, 37, 12</td>
<td>No</td>
</tr>
</tbody>
</table>
Different Decisions = Different Levels of Care . . . i.e. Different Levels of Exposure Risk
Data Interpretation Example

- Employee performs a job 100 times per year
- If you collected personal samples on the employee all 100 times, how many times is it acceptable for exposures to exceed the Occupational Exposure Limit (OEL) without a respirator?
  - 0 samples?
  - 1 sample?
  - 5 samples?
  - 10 samples?
  - 25 samples?
  - 50 samples?
Why the Inconsistencies?

- Variable Definitions of Acceptable
How much assurance?

- 100% Sure?
- 95%?
- 90%?
- 85%?
- 50%?
Why the Inconsistencies?

- Variable Definitions of Acceptable
- Uncertainty
Exposure Judgments

Inputs
- Basic Characterization Info
- Training
- Experience

Outputs
- Exposure Judgment
  - Exposure Estimate
  - Hazard Estimate
  - Uncertainty Estimate
  - Acceptability Estimate

Black Box
“OH Professional Judgment”
Basic Characterization

Exposure Assessment

Acceptable Exposure
Uncertain
Unacceptable Exposure

Control
Further Information Gathering

Reassessment
# The AIHA “Exposure Control Banding” Model

<table>
<thead>
<tr>
<th>Control Zone Description</th>
<th>AIHA Recommended Statistical Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly-controlled (HC)</td>
<td>(X_{0.95} \leq 0.10 \text{ OEL})</td>
</tr>
<tr>
<td>Well-controlled (WC)</td>
<td>(0.10 \text{ OEL} &lt; X_{0.95} \leq 0.5 \text{ OEL})</td>
</tr>
<tr>
<td>Controlled (C)</td>
<td>(0.5 \text{ OEL} &lt; X_{0.95} \leq \text{ OEL})</td>
</tr>
<tr>
<td>Poorly controlled (PC)</td>
<td>(\text{OEL} &lt; X_{0.95})</td>
</tr>
</tbody>
</table>
Data Interpretation Exercise
Class Work: DIT
Add the probability of the 95th Percentile being in each of the 4 categories.

There must be only ONE highest category

The total probability for all 4 categories must be equal to 100%
## Exposure Ratings

<table>
<thead>
<tr>
<th>Exposure Rating</th>
<th>Cutoff (%OEL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$X_{0.95} \leq 10%$</td>
</tr>
<tr>
<td>2</td>
<td>$10% &lt; X_{0.95} \leq 50%$</td>
</tr>
<tr>
<td>3</td>
<td>$50% &lt; X_{0.95} \leq 100%$</td>
</tr>
<tr>
<td>4</td>
<td>$X_{0.95} &gt; 100%$</td>
</tr>
</tbody>
</table>
Probability Chart for 95%tile Exposure Judgements

Example - "There is a 45% probability that the 95%tile falls between 10% & 50% of the OEL."
Example of filling out the DIT

<table>
<thead>
<tr>
<th>Categories</th>
<th>Dataset #1 - Probability of 95\textsuperscript{th} Percentile in Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;10%</td>
<td>30</td>
</tr>
<tr>
<td>10-50%</td>
<td>45</td>
</tr>
<tr>
<td>50-100%</td>
<td>15</td>
</tr>
<tr>
<td>&gt;100%</td>
<td>10</td>
</tr>
</tbody>
</table>
**Data Interpretation Test (DIT) #7**

Enter Your Number

<table>
<thead>
<tr>
<th>OEL for all Data Sets</th>
<th>Sample Data Set #1</th>
<th>Sample Data Set #2</th>
<th>Sample Data Set #3</th>
<th>Sample Data Set #4</th>
<th>Sample Data Set #5</th>
<th>Sample Data Set #6</th>
<th>Sample Data Set #7</th>
<th>Sample Data Set #8</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>18</td>
<td>9</td>
<td>16</td>
<td>71</td>
<td>6</td>
<td>19</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>31</td>
<td>19</td>
<td>4</td>
<td>38</td>
<td>23</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>107</td>
<td>1</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Make your judgments on the above Statistics Test Data in the following columns

<table>
<thead>
<tr>
<th>1-10% OEL</th>
<th>Data Set #1</th>
<th>Data Set #2</th>
<th>Data Set #3</th>
<th>Data Set #4</th>
<th>Data Set #5</th>
<th>Data Set #6</th>
<th>Data Set #7</th>
<th>Data Set #8</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-50% OEL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-100% OEL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;100% OEL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Check</td>
<td>100?</td>
<td>100?</td>
<td>100?</td>
<td>100?</td>
<td>100?</td>
<td>100?</td>
<td>100?</td>
<td>100?</td>
</tr>
</tbody>
</table>

Have you ever taken this statistical test before? **Yes** **No**

If yes, how many times & when?

**Instructions**

- Enter your name at the top.
- Review each data set and document the probabilities of where the 95th%tile falls.
- Make sure that one category has the highest percentage.
- Do not enter values less than 1 in any field (no zeros!).
- Check to see that each Data Set Column adds to 100%.

Please list any specific comments regarding this DIT.
Review of Lognormal Statistics and analyzing small data sets

Slides courtesy of
Paul Hewett Ph.D. CIH
Exposure Assessment Solutions, Inc.
Review of IH Statistics

I. Lognormal distribution
II. Sample 95\textsuperscript{th} percentile
III. UCL for the sample 95\textsuperscript{th} percentile
IV. Rules-of-thumb for “Eyeballing” Exposure Data
I. Lognormal Distribution – Example
Airborne exposures to inorganic lead

source: Cope et al. AIHAJ 40:372-379, 1979
### Parameters vs. Statistics

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-calculated using all elements of the population</td>
<td>-calculated from a sample of n elements randomly selected</td>
</tr>
<tr>
<td>Population Geometric Mean</td>
<td>Sample Geometric Mean</td>
</tr>
<tr>
<td>GM</td>
<td>gm</td>
</tr>
<tr>
<td>Population Geometric Standard Deviation</td>
<td>Sample Geometric Standard Deviation</td>
</tr>
<tr>
<td>GSD</td>
<td>gsd</td>
</tr>
</tbody>
</table>

Copyright 2007 Exposure Assessment Solutions, Inc.
Lognormal distribution PDF

gm = 1.06 mg/m³
gsd = 1.83
Sample geometric mean (gm) & geometric standard deviation (gsd)

\[
let \ y = \ln(x) \\

gm = \exp\left(\frac{\sum y_i}{n}\right) \\

gsd = \exp\sqrt{\frac{\sum(y_i - \bar{y})^2}{n-1}}
\]
Example: Welding fume data - estimate GM and GSD

<table>
<thead>
<tr>
<th>Case</th>
<th>$x_i$ (mg/ m$^3$)</th>
<th>$y_i = \ln(x_i)$</th>
<th>$(y_i - \bar{y})^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.84</td>
<td>-0.1744</td>
<td>0.055877</td>
</tr>
<tr>
<td>2</td>
<td>0.98</td>
<td>-0.0202</td>
<td>0.006762</td>
</tr>
<tr>
<td>3</td>
<td>0.42</td>
<td>-0.8675</td>
<td>0.864025</td>
</tr>
<tr>
<td>4</td>
<td>1.16</td>
<td>0.1484</td>
<td>0.007463</td>
</tr>
<tr>
<td>5</td>
<td>1.36</td>
<td>0.3075</td>
<td>0.060248</td>
</tr>
<tr>
<td>6</td>
<td>2.66</td>
<td>0.9783</td>
<td>0.839600</td>
</tr>
</tbody>
</table>

**Sum =**

\[ \sum_{i=1}^{6} (y_i - \bar{y})^2 = 1.833976 \]

\[ \bar{y} = 0.0620 \]

\[ \text{gm} = 1.06 \]

\[ \text{gsd} = 1.83 \]
Example: Welding fume data - estimate GM and GSD

\[ gm = \exp\left( \frac{0.3722}{6} \right) = 1.06 \text{ mg/m}^3 \]

\[ gsd = \exp\left( \frac{1.833976}{6 - 1} \right) = 1.83 \]
Example: Welding fume data - estimate \( \mu \) and \( \sigma \)

<table>
<thead>
<tr>
<th>Case</th>
<th>( x_i ) (mg/ m(^3))</th>
<th>( (x_i - \bar{x})^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.84</td>
<td>0.157344</td>
</tr>
<tr>
<td>2</td>
<td>0.98</td>
<td>0.065878</td>
</tr>
<tr>
<td>3</td>
<td>0.42</td>
<td>0.666944</td>
</tr>
<tr>
<td>4</td>
<td>1.16</td>
<td>0.005878</td>
</tr>
<tr>
<td>5</td>
<td>1.36</td>
<td>0.015211</td>
</tr>
<tr>
<td>6</td>
<td>2.66</td>
<td>2.025878</td>
</tr>
<tr>
<td>Sum</td>
<td>7.42</td>
<td>2.937133</td>
</tr>
<tr>
<td>( \bar{x} )</td>
<td>1.24</td>
<td></td>
</tr>
<tr>
<td>( sd )</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>
Example: Welding fume data - estimate $\mu$ and $\sigma$

$$
\bar{x} = \frac{7.42}{6} = 1.24 \text{ mg/m}^3
$$

$$
\sigma = \sqrt{\frac{2.9371636}{6-1}} = 0.77 \text{ mg/m}^3
$$
Sample 95\textsuperscript{th} Percentile Exposure

- The focus is on the upper tail of the exposure profile.
- The sample 95\textsuperscript{th} percentile can be considered a “decision statistic”.
- The (usual) goal is to determine which category the 95\textsuperscript{th} Percentile most likely falls.
- It is used to assist in reaching a decision that the exposure profile is
  - “Controlled” or “Acceptable”
  - “Unacceptable”
  - or falls in a “Control Category”

Copyright 2007 Exposure Assessment Solutions, Inc.
95th Percentile interpretation of TWA OELs

- ACGIH
95th Percentile interpretation of TWA OELs

- AIHA 1991 and 1998 guidance
  - Employer should maintain true group or individual upper percentile exposure < TWA OEL
  - “Similar Exposure Group” 95th percentile exposure < TWA OEL

95th Percentile interpretation of TWA OELs

- NIOSH guidance
  - Employer should 95% confident that 95% of the exposures are ≤ the TWA PEL

- OSHA
  - Measured TWA exposures should “rarely” exceed the TWA PEL (preamble to the benzene PEL, 1987)
95th Percentile interpretation of TWA OELs

EU
Example

A sample of six full-shift TWA welding fume measurements resulted in the following statistics:

- (sample) geometric mean is 1.06 mg/m$^3$
- (sample) geometric standard deviation is 1.83

What is the point estimate (i.e., best estimate) of the true 95th percentile?
90\textsuperscript{th}, 95\textsuperscript{th}, and 99\textsuperscript{th} Percentiles

\[ \bar{y} = \ln(\bar{g}) \]

\[ \bar{s}_y = \ln(\bar{d}) \]

\[ \hat{X}_p = \exp(\bar{y} + Z_p \cdot s_y) \]

\[ \hat{X}_{0.90} = \exp(\bar{y} + 1.282 \cdot s_y) \]

\[ \hat{X}_{0.95} = \exp(\bar{y} + 1.645 \cdot s_y) \]

\[ \hat{X}_{0.99} = \exp(\bar{y} + 2.327 \cdot s_y) \]
\[ X_{0.95} = \exp(\bar{y} + 1.645 \cdot s_y) \]

\[ = \exp(0.0620 + 1.645 \cdot 0.6043) \]

\[ = 2.88 \; \text{mg/m}^3 \]
\[ \hat{X}_p = \text{gm} \cdot \text{gsd}^{1.645} \]

Alternative upper percentile formula

\[ \hat{X}_{0.95} = \text{gm} \cdot \text{gsd}^{1.645} \]

\[ = 1.06 \cdot 1.83^{1.645} = 2.88 \text{ mg/m}^3 \]
Focus on Upper Tail

\[ gm = 1.06 \text{ mg/m}^3 \]
\[ gsd = 1.83 \]

point estimate of the 95th percentile
III. Upper Confidence Limit (UCL) for the Sample 95\textsuperscript{th} Percentile

- Calculate confidence intervals around estimates of …
  - upper percentile (normal & lognormal)
- Confidence intervals are used to …
  - express uncertainty
  - test hypotheses:
    - to determine our confidence level that the SEG is in compliance with an OEL
    - to determine our confidence level that the true 95\textsuperscript{th} percentile exposure is within a specific exposure control category
For single shift, TWA exposure limits (TWA OELs) …

- focus on the upper tail of the distribution
- e.g., 95th percentile exposure
Upper Percentile (e.g., 95th percentile)

- **Concept**
  - Calculate the 95% upper confidence interval for the 95th percentile statistic (upper tolerance limit)

- **Application**
  - 95%UCL can be used to test the following hypotheses:
    - $H_0$: 95th percentile $> \text{OEL}$
    - $H_a$: 95th percentile $< \text{OEL}$

- **Interpretation**
  - If the 95%UCL is less than the OEL, then we can say that we are at least 95% confident that the true 95th percentile is less than the OEL.
95% UCL for the 95\textsuperscript{th} Percentile

- **Procedure:**
  - Calculate the gm and gsd
  - Using \( n \), read the UCL \( K \)-value from the appropriate table
    - \( \gamma \) = confidence level, e.g., 0.95
    - \( p \) = proportion, e.g., 0.95
    - \( n \) = sample size
  - Using \( \hat{gm}, \hat{gsd}, \) and \( k \), calculate the 95% UCL
    - \( y = \ln(\,gm\,) \)
    - \( s_y = \ln(\,gsd\,) \)

\[
95\%UCL(\hat{X}_{0.95}) = \exp(\,\bar{y} + K_{\gamma,p,n} s_y)\]

Copyright 2007 Exposure Assessment Solutions, Inc.
### TABLE VII.3 — Factors for One-Sided Tolerance Limits

<table>
<thead>
<tr>
<th>( \gamma = 0.95 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
</tbody>
</table>
\[ 95\% UCL(\hat{X}_{0.95}) = \exp(\bar{y} + K_{\gamma, p, n} \cdot s_y) \]
\[ = \exp(\bar{y} + K_{0.95, 0.95, 6} \cdot s_y) \]
\[ = \exp(0.0620 + 3.707 \cdot 0.6043) \]
\[ = 10.00 \text{ mg/m}^3 \]
IV. Rule-of-thumb for “Eyeballing” Exposure Data

Given:

- $G = \text{median}$
- $X_p = G \times D^{Z_p}$  (e.g., $X_{0.95} = G \times D^{1.645}$)

… a Rule-of-thumb, or guideline, can be devised for quickly estimating from limited data the range in which the true 95th percentile might lie.
<table>
<thead>
<tr>
<th>GSD</th>
<th>Multiple of GM (median)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$X_p = 95^{th}$ percentile</td>
</tr>
<tr>
<td></td>
<td>$Z_p = 1.645$</td>
</tr>
<tr>
<td>1.5</td>
<td>1.95</td>
</tr>
<tr>
<td>2.0</td>
<td>3.13</td>
</tr>
<tr>
<td>2.5</td>
<td>4.51</td>
</tr>
<tr>
<td>3.0</td>
<td>6.09</td>
</tr>
</tbody>
</table>
R.O.T. for Estimating the 95\textsuperscript{th} Percentile

1. If $n$ is small (i.e., $<6$) and one or more measurements $>\text{OEL}$, then decision = Category 4.

2. Estimate the median and use it as a surrogate of the sample GM:
   - Sort the data
   - If $n$ is odd the median is the middle value.
   - If $n$ is even the median is the average of two middle values.

3. Multiply the median by 2, 4, and 6
   - The results comprise an \textit{approximate} low, middle, and high estimate of $X_{0.95}$. 

Copyright 2007 Exposure Assessment Solutions, Inc.
Rule-of-thumb Workshop  
(assume OEL=100)

a. \( X = \{5\} \)
b. \( X = \{68\} \)
c. \( X = \{7, 34, 57\} \)
d. \( X = \{1, 1, 2, 5\} \)
e. \( X = \{4, 5, 8, 23\} \)
f. \( X = \{0.3, 1, 2, 3, 4, 22\} \)
g. \( X = \{10, 10, 10, 20, 50, 105\} \)
h. \( X = \{7, 10, 16, 21, 45, 53\} \)

For each dataset, determine the appropriate Exposure Category – 1, 2, 3, or 4 – using the above Rule-of-thumb.
Chart of 100 Industrial Hygiene Air samples - Logarithmic Data
Chart of 100 Industrial Hygiene Air samples - Logarithmic Data
Problems with judging or estimating 95% tiles

- Limited data for many jobs or tasks
- Very large statistical confidence intervals with small data sets
- Censored Data (Below LOD)
- Log data can be difficult to judge
- Difficult to Communicate
Is the exposure represented by these samples acceptable?

Ethanol OEL = 1000 ppm

Monitoring Results:

<table>
<thead>
<tr>
<th>Concentration (ppm)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>215</td>
<td>ppm</td>
</tr>
<tr>
<td>52</td>
<td>ppm</td>
</tr>
<tr>
<td>395</td>
<td>ppm</td>
</tr>
<tr>
<td>700</td>
<td>ppm</td>
</tr>
<tr>
<td>75</td>
<td>ppm</td>
</tr>
</tbody>
</table>

Traditional IH Statistics:

- GM = 188
- GSD = 3
- 95th percentile = 1140 ppm
- UTL_{95\%,95\%} = 18,700 ppm
Is the exposure represented by these samples acceptable?

Ethanol OEL = 1000 ppm

Monitoring Results:
215 ppm
52 ppm
395 ppm
700 ppm
75 ppm
Bayesian Decision Analysis (BDA)

- An adjunct or alternative to the calculation and interpretation of traditional statistics.

- The goal of BDA is to estimate the probability that the true exposure profile falls into a particular category, or Exposure Rating.
Straightforward Interpretation: Bayesian Likelihood Distribution

Likelihood that 95%ile falls into indicated Exposure Rating Category

<table>
<thead>
<tr>
<th>Probability</th>
<th>&lt;1% OEL</th>
<th>&lt;10% OEL</th>
<th>10 – 50%</th>
<th>50 – 100%</th>
<th>&gt;100% OEL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0.087</td>
<td>0.4</td>
<td>0.513</td>
</tr>
</tbody>
</table>

Exposure Rating Category:
- <1% OEL
- <10% OEL
- 10 – 50%
- 50 – 100%
- >100% OEL
Much easier to communicate!
AIHA Model: Inherently Bayesian
AIHA Exposure Assessment Flow Diagram

Start

Basic Characterization

Exposure Assessment

Acceptable Exposure

Uncertain

Unacceptable Exposure

Control

Further Information Gathering

Reassessment
AIHA EA Strategy

Define Exposure Using All Available Information

Conditions

Qualitative Modeling

EA Tools

Monitoring

Exposure Profile
Initial Assessment

Heavy Emphasis on Professional Judgment or Modeling

Conditions

Qualitative Modeling

Monitoring

EA Tools

Exposure Profile

Start

Basic Characterization

Exposure Assessment

Acceptable Exposure

Uncertain Exposure

Unacceptable Exposure

Control

Further Information Gathering

Reassessment
Validated Assessment

Heavy Emphasis on Monitoring Data

Conditions

Qualitative Modeling  EA Tools Monitoring

Exposure Profile
AIHA EA Strategy:

Define Exposure **Using All Available Information**

**Conditions**

- Qualitative Modeling
- Monitoring

**Exposure Profile**

**EA Tools**
Qualitative Assessment or Validated Model

- Prior probability:
  - 0.8
  - 0.5

- Decision probabilities:
  - 0.6
  - 0.4
  - 0.2

- Monitoring:
  - Exposure Rating
  - Start Basic Characterization
  - Exposure Assessment
  - Acceptable Exposure
  - Uncertain Exposure
  - Uncertain Exposure
  - Unacceptable Exposure
  - Control
  - Reassessment
  - Further Information Gathering

- Graph showing decision probability distribution for exposure ratings from 0 to 4.
Qualitative Modeling

Monitoring

Qualitative Assessment or Validated Model

Monitoring Results
Industrial Hygienists Are Bayesian Thinkers!

Qualitative Assessment or Validated Model

Monitoring Results

Integrated Exposure Assessment
### Example: Exposure Rating Category Follow-up

<table>
<thead>
<tr>
<th>Exposure Control Category**</th>
<th>Recommended Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>0  (&lt;1% of OEL)</td>
<td>No action</td>
</tr>
<tr>
<td>1  (&lt;10% of OEL)</td>
<td>general HazCom</td>
</tr>
<tr>
<td>2  (10-50% of OEL)</td>
<td>+ chemical specific HazCom</td>
</tr>
<tr>
<td>3  (50-100% of OEL)</td>
<td>+ exposure surveillance, medical surveillance, work practices</td>
</tr>
<tr>
<td>4  (&gt;100% of OEL)</td>
<td>+ respirators &amp; engineering controls, work practice controls</td>
</tr>
<tr>
<td>5  (Multiples of OEL; e.g., based on respirator APFs)</td>
<td>+ immediate engineering controls or process shutdown, validate respirator selection</td>
</tr>
</tbody>
</table>

** - Decision statistic = 95th percentile
An Example Using the AIHA Model
Example: Exposure Estimate

Simple Model:

\[ C = \frac{G}{Q} \]

Worst Case

\[ C = \frac{65 \text{ mg/hour}}{3.6 \text{ m}^3/\text{hour}} = 18 \text{ mg/m}^3 \]

Best Case

\[ C = \frac{35 \text{ mg/hour}}{540 \text{ m}^3/\text{hour}} = 0.065 \text{ mg/m}^3 \]
Uncertainty and Acceptability

Concentration (mg/M^3)

18
Simple Model

0.065
Example: Exposure Estimate

Statistical Modeling: Monte Carlo
Uncertainty Analysis

Agent “X”
G = steady generation rate (mg/hour)
   35 to 65 mg/hour
Q = steady ventilation rate (m$^3$/hour)
   3.6 to 540 m$^3$/hour

\[ C = \frac{G}{Q} \]

Forecast: Concentration
10,000 Trials

Mean = 0.46
Certainty is 95.30% from 0.00 to 1.75 mg/m$^3$
Uncertainty and Acceptability

Monte Carlo

Simple Model

Concentration (mg/M$^3$)

0.065 1.75 0.22

18
Uncertainty and Acceptability

Which To Choose? Acceptable?
Uncertainty and Acceptability

Any or all of these can be used to build the Bayesian Qualitative Model

Which To Choose? Acceptable?
Uncertainty and Acceptability

Monte Carlo

Simple Model

Concentration (mg/M³)

0.065
1.75
0.22

18

OEL = 10
AIHA EA Strategy:
Define Exposure Using All Available Information

Add Monitoring Data . . .
Validate Initial Judgments
Example: Exposure Estimate

Monitoring Results:

<table>
<thead>
<tr>
<th>Concentration (mg/M³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
</tr>
<tr>
<td>0.14</td>
</tr>
<tr>
<td>0.21</td>
</tr>
<tr>
<td>0.37</td>
</tr>
<tr>
<td>0.78</td>
</tr>
</tbody>
</table>

Agent “X”

G = steady generation rate (mg/hour)
   35 to 65 mg/hour

Q = steady ventilation rate (m³/hour)
   3.6 to 540 m³/hour

95%ile

UTL_{95\%,95\%} = 16 mg/M³
Example: Exposure Estimate

Agent "X"
G = steady generation rate (mg/hour)
  35 to 65 mg/hour
Q = steady ventilation rate (m$^3$/hour)
  3.6 to 540 m$^3$/hour

Monitoring Results:

<table>
<thead>
<tr>
<th>Concentration (mg/M$^3$)</th>
<th>Exposure Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1</td>
</tr>
<tr>
<td>0.14</td>
<td>1</td>
</tr>
<tr>
<td>0.21</td>
<td>2</td>
</tr>
<tr>
<td>0.37</td>
<td>3</td>
</tr>
<tr>
<td>0.78</td>
<td>4</td>
</tr>
</tbody>
</table>

Likelihood

Decision Probability

Exposure Rating
Qualitative Assessment or Validated Model

Prior Probability

Decision Probability

Exposure Rating

Monte Carlo

COSH H Ess.

OEL = 1

Monitoring Results

Simple Model

0.05 mg/M³
0.14 mg/M³
0.21 mg/M³
0.37 mg/M³
0.78 mg/M³
## Integrated Exposure Assessment Result

Leads to Control Recommendations

### Exposure Control Category

<table>
<thead>
<tr>
<th>Category</th>
<th>Recommended Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (&lt;1% of OEL)</td>
<td>No action</td>
</tr>
<tr>
<td>1 (&lt;10% of OEL)</td>
<td>General Haz Com</td>
</tr>
<tr>
<td>2 (10-50% of OEL)</td>
<td>+ Chemical specific Haz Com</td>
</tr>
<tr>
<td>3 (50-100% of OEL)</td>
<td>+ Medical surveillance, work practices</td>
</tr>
<tr>
<td>4 (&gt;100% of OEL)</td>
<td>+ Respirators &amp; engineering controls, work practice controls</td>
</tr>
<tr>
<td>Multiples of OEL (i.e., based on respirator APFs)</td>
<td>+ Immediate Engineering Controls or Process Shut Down, Validate Acceptable Respirator</td>
</tr>
</tbody>
</table>
Example Survey

• OEL = 1 ppm
• During a baseline/initial exposure assessment, an IH collected the following full-shift measurements from an SEG:
  – 0.20, 0.05, & 0.10 ppm
• n = 3; gm = 0.10; gsd = 2.00
• The sample 95\textsuperscript{th} percentile was 0.31 ppm
• but with a 95\%UCL of 20 ppm
When n is small, confidence intervals are often extremely broad.

- \( X = \{0.20, 0.05, 0.10 \text{ ppm}\} \)
- \( n = 3 \)

- \( \text{gm} = 0.1 \text{ ppm} \quad 90\%\text{CI}(0.03, 0.32) \)
- \( \text{gsd} = 2.0 \quad 90\%\text{CI}(1.5, 21) \)
- \( X_{0.95} = 0.31 \text{ ppm} \quad 90\%\text{CI}(0.16, 20) \)
Example Survey (cont’d)

• The point estimate of the 95\textsuperscript{th} percentile is < 50\% of the limit.
• Exposures appear to be a Category 2 exposure.
• However, the 95\%UCL(X_{0.95}) is considerably greater than the OEL.
• What would you do?
  – Make a decision?
  – Collect more data?
Example (cont’d)

• Our IH concludes:
  – This operation is well-controlled with just the existing dilution ventilation.
  – Although the 95%UCLs were excessive, our IH took into account his extensive past experience with this type of operation.

• His recommendations:
  – Further sampling is not necessary.
  – Routine surveillance samples should be collected using the established schedule for well-controlled operations.

• Is such a decision making process a Bayesian Decision Analysis?
Exposures appear to be a Category 2 exposure.

0.20 ppm
0.05 ppm
0.10 ppm

X0.95 = 0.31 ppm
90%CI (0.16, 20)

Our IH concludes:
This operation is well-controlled with just the existing dilution ventilation.
Bayes’ Theorem – The Foundation of Bayesian Statistics

\[
P(\text{Pop}_i|\text{data}) = \frac{P(\text{data}|\text{Pop}_i) \cdot P(\text{Pop}_i)}{\sum_i [P(\text{data}|\text{Pop}_i) \cdot P(\text{Pop}_i)]}
\]

Correction Factor

Posterior  Likelihood  Prior
Who gave you the ugly tie?

- At your birthday party you receive a truly ugly tie. The wrapping was plain, with no label.
- Who gave you the tie?
- The choices are the stingy aunt and the weird uncle.
- Considering the two, the chances that your aunt or uncle would bring a gift are 1 in 4 and 3 in 4, respectively.
- The probability of your aunt giving you an ugly tie is low; for example, 1 in 10.
- The probability of your uncle giving you an ugly tie is high; for example, 1 in 2.
Given these two choices – and the Prior and Likelihood estimates/guesses - there is nearly a 94% probability that the your uncle gave you the tie.
III. Bayes’ Theorem Applied to Exposure Profiles

\[ P(\ln G_i, \ln D_i \mid data) = \frac{P(data \mid \ln G_i, \ln D_i) \cdot P(\ln G_i, \ln D_i)}{\sum_{i=1}^{k} [P(data \mid \ln G_i, \ln D_i) \cdot P(\ln G_i, \ln D_i)]} \]

Equation 1
Likelihood Function

• The relative probability of the data, given an exposure profile is calculated using the likelihood function (y=ln(x)):

\[ P(\text{data} \mid \ln G_i, \ln D_i) \propto \prod_{j=1}^{n} pdf\left(y_j \mid \ln G_i, \ln D_i\right) \]

Equation 2

\[ pdf(y \mid \ln G_i, \ln D_i) = \frac{1}{\ln D_i \sqrt{2\pi}} \cdot \exp\left(\frac{-\left(y - \ln G_i\right)^2}{2(ln D_i)^2}\right) \]
Simple Example – Two Exposure Profiles

• Say we are interested in determining which of two exposure profiles is most likely.

• Exposure Profile A
  – GM = 0.15 ppm
  – GSD = 2

• Exposure Profile B
  – GM = 0.25 ppm
  – GSD = 2
Prior

• Let us assign *a priori* probabilities of Exposure Profile A and B:
  – Prob(A) = 0.7
  – Prob(B) = 0.3

• Then collect some data:
  \( x = \{0.20, 0.05, 0.10\} \text{ ppm} \)
• Calculate the PDF values for each exposure profile:

• Calculate the product of the PDF values for each exposure profile.
Likelihood Function

• To display the Likelihood Probabilities the Likelihood Function must be normalized:

\[
P(data \mid \ln G_i, \ln D_i) = \frac{\prod_{j=1}^{n} pdf(y_j \mid \ln G_i, \ln D_i)}{\sum_{i=1}^{k} \left[ \prod_{j=1}^{n} pdf(y_j \mid \ln G_i, \ln D_i) \right]}
\]

• \( n \) = number of measurements
• \( k \) = number of exposure profiles

Equation 2
Likelihood Function (cont’d)

- Display the Likelihood Decision probabilities in the Likelihood Decision Chart:
Posterior Function

• Combine the Prior and Likelihood functions using Bayes’ Equation:

\[
P(\ln G_i, \ln D_i | data) = \frac{P(data | \ln G_i, \ln D_i) \cdot P(\ln G_i, \ln D_i)}{\sum_{i=1}^{k} [P(data | \ln G_i, \ln D_i) \cdot P(\ln G_i, \ln D_i)]}
\]
Posterior Function

- Display the Posterior Decision probabilities in the Posterior Decision Chart:
• The original Bayes’ Theorem directly applies to discrete choices.
  – e.g., Exposure Profiles A vs. B

• We are not interested in distinguishing between just two exposure profiles.

• Instead, we are interested in distinguishing between five *populations* of exposure profiles:
  – Exposure Zones 0, 1, 2, 3, and 4
Rating Exposure Control Using Bayesian Decision Analysis

Paul Hewett,¹ Perry Logan,² John Mulhausen,² Gurumurthy Ramachandran,³ and Sudipto Banerjee³
### Exposure Ratings – A “rating zone” represents a population of exposure profiles

<table>
<thead>
<tr>
<th>Exposure Rating</th>
<th>Cutoff (%OEL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$X_{0.95} \leq 1%$</td>
</tr>
<tr>
<td>1</td>
<td>$1% &lt; X_{0.95} \leq 10%$</td>
</tr>
<tr>
<td>2</td>
<td>$10% &lt; X_{0.95} \leq 50%$</td>
</tr>
<tr>
<td>3</td>
<td>$50% &lt; X_{0.95} \leq 100%$</td>
</tr>
<tr>
<td>4</td>
<td>$X_{0.95} &gt; 100%$</td>
</tr>
</tbody>
</table>
Exposure Ratings translated into parameter space for OEL=1ppm
Exposure Ratings translated into parameter space for OEL=1 ppm
Prior Decision Distribution

• Categorical
  – Assign an \textit{a priori} probability to each Exposure Rating zone

![Prior Decision Distribution Graph]

- Prior
  - Exposure Rating
  - Decision Probability

<table>
<thead>
<tr>
<th>Exposure Rating</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
</tr>
<tr>
<td>4</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Prior decision function (i.e., prior decision distribution spread across parameter space)

Figure 2: The Prior Decision Distribution expressed as a Prior Function spread throughout parameter space.
Prior decision function (i.e., prior decision distribution spread across parameter space)
Example *Likelihood* Decision Distribution for $x = \{0.20, 0.05, 0.10\}$
Likelihood function for $x=\{0.20, 0.05, 0.10\}$

Figure 3: The Likelihood Function calculated using Equation 5 and the example dataset: $x=\{0.20, 0.05, 0.10\}$.
Figure 4: The Posterior Function (Equation 4), which is the product of the Prior and Likelihood Functions.
BDA

Prior

Decision Probability

Exposure Rating

Likelihood

Decision Probability

Exposure Rating

Posterior

Decision Probability

Exposure Rating
Decision Charts

- OEL = 1 ppm
- n = 3
- x = {0.20, 0.05, 0.10} ppm

- Here we used a uniform prior (also called Flat or Non-informative prior).
Decision Charts

- OEL=1 ppm
- $n = 3$
- $x = \{0.20, 0.05, 0.10\}$ ppm

- Here we used an informative prior.
Scenario #1 – Process Operator #1

• Process Operator #1 is responsible for the following tasks
  – Opening a valve that directly charges xylene into the process mixer
  – Manually charging solids into the process mixer (75 pounds once per hour)
  – Collecting multiple quality samples once each hour through manhole
  – No previous personal air samples available

• We’ve collected some full shift air samples for xylene, now let’s do some BDA!
  – 13 ppm, 26 ppm, 18 ppm
Enter information and sampling data & Press “Calculate All”

Let's focus on the Likelihood (ie. No prior knowledge).
How do we interpret this?

- The output is in probability
- “We have a __% probability that Process Operator #1 requires additional exposure controls”

- Is that above the acceptable / unacceptable threshold?
Compare BDA vs. traditional statistics…

• “We have a ___% probability that Process Operator #1 requires additional exposure controls”

• “The population 95\textsuperscript{th} percentile point estimate is 32 with an upper confidence limit (95%) of 260”
How would you interpret this?
More examples…

“less than __ % probability of…” or “greater than __ % probability of…”
More Examples...

• “given our sampling data, we have a greater than 95% probability that exposures are acceptable…”
• “greater than 27% probability that exposures are unacceptable…”
• “less than 10% probability that exposures exceed our medical surveillance triggers…”
• “greater than 95% probability that exposures require immediate exposure controls…”
Limited Data (OEL = 1 ppm)

- 0.2 ppm
- 0.05 ppm
- 0.1 ppm
Limited Data (OEL = 5 mg/M³)

- Dataset24 - Welding Fumes.xls
- Measurements collected from an SEG on a single day at a frame manufacturing facility in 1987.

1.63 mg/ M³
4.28 mg/ M³
2.04 mg/ M³
2.32 mg/ M³
2.02 mg/ M³
6.04 mg/ M³
Limited Data (OEL = 5 mg/M³)

- Measurements collected from an SEG on a single day at a frame manufacturing facility in 1987.

<table>
<thead>
<tr>
<th>Exposure Rating</th>
<th>Decision Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.84 mg/M³</td>
</tr>
<tr>
<td>1</td>
<td>0.98 mg/M³</td>
</tr>
<tr>
<td>2</td>
<td>0.42 mg/M³</td>
</tr>
<tr>
<td>3</td>
<td>1.16 mg/M³</td>
</tr>
<tr>
<td>4</td>
<td>1.36 mg/M³</td>
</tr>
<tr>
<td>5</td>
<td>2.66 mg/M³</td>
</tr>
<tr>
<td>6</td>
<td>0.48 mg/M³</td>
</tr>
<tr>
<td>7</td>
<td>0.92 mg/M³</td>
</tr>
<tr>
<td>8</td>
<td>1.32 mg/M³</td>
</tr>
<tr>
<td>9</td>
<td>1.72 mg/M³</td>
</tr>
<tr>
<td>10</td>
<td>2.12 mg/M³</td>
</tr>
<tr>
<td>11</td>
<td>2.52 mg/M³</td>
</tr>
<tr>
<td>12</td>
<td>2.92 mg/M³</td>
</tr>
<tr>
<td>13</td>
<td>3.32 mg/M³</td>
</tr>
<tr>
<td>14</td>
<td>3.72 mg/M³</td>
</tr>
<tr>
<td>15</td>
<td>4.12 mg/M³</td>
</tr>
<tr>
<td>16</td>
<td>4.52 mg/M³</td>
</tr>
<tr>
<td>17</td>
<td>4.92 mg/M³</td>
</tr>
<tr>
<td>18</td>
<td>5.32 mg/M³</td>
</tr>
<tr>
<td>19</td>
<td>5.72 mg/M³</td>
</tr>
<tr>
<td>20</td>
<td>6.12 mg/M³</td>
</tr>
<tr>
<td>21</td>
<td>6.52 mg/M³</td>
</tr>
<tr>
<td>22</td>
<td>6.92 mg/M³</td>
</tr>
<tr>
<td>23</td>
<td>7.32 mg/M³</td>
</tr>
<tr>
<td>24</td>
<td>7.72 mg/M³</td>
</tr>
<tr>
<td>25</td>
<td>8.12 mg/M³</td>
</tr>
<tr>
<td>26</td>
<td>8.52 mg/M³</td>
</tr>
<tr>
<td>27</td>
<td>8.92 mg/M³</td>
</tr>
<tr>
<td>28</td>
<td>9.32 mg/M³</td>
</tr>
<tr>
<td>29</td>
<td>9.72 mg/M³</td>
</tr>
<tr>
<td>30</td>
<td>10.12 mg/M³</td>
</tr>
</tbody>
</table>

![Graph showing likelihood of exposure ratings](image-url)
Single measurement scenario

- Let OEL = 100 ppm

5 ppm

50 ppm

99 ppm

150 ppm
Large Datasets (OEL = 0.05 mg/M$^3$)

- N=15
- Inorganic lead

\[
\begin{array}{ll}
0.012 & 0.0081 \\
0.0109 & 0.012 \\
0.0086 & 0.0081 \\
0.0382 & 0.0194 \\
0.0073 & 0.029 \\
0.0138 & 0.0183 \\
0.0108 & 0.0306 \\
0.0103 & \\
\end{array}
\]
BDA Options: change exposure category cutoffs
Change dimensions of the Parameter Space: $\text{GM}_{\text{min}}$, $\text{GM}_{\text{max}}$, $\text{GSD}_{\text{min}}$, and $\text{GSD}_{\text{max}}$
Process Operator #1

Storage Tank

Reactor

Process Operator #2

Process Engineer

Lets focus on Process Operator #2
Scenario #2 – Process Operator #2

• Process Operator #2 is responsible for the following tasks
  – Filling products into drums 4 times per shift (a new drum local exhaust ventilation is available)
  – Manually changing filter media once per shift and periodically using xylene solution to clean filtering equipment as needed to remove plugs
  – Collecting 6 – 3 oz quality samples on each batch.

• We’ve collected some full shift air samples for xylene, now lets do some BDA!
Enter information and sampling data & Press “Calculate All”

Let's focus on the Likelihood (ie. No prior knowledge).

Take a good look at the data!!! Any comments?
How do we interpret this?

- “We have less than a ___% probability that Process Operator #2 requires respiratory protection”
- Is it above the acceptable / unacceptable threshold?
- Are there any other observations? Let's take a closer look at the data… (1 ppm, 65 ppm, 0.5 ppm)
What about our sample GSD?

What is the impact on the analysis???

Let's take a journey into our universe...

<table>
<thead>
<tr>
<th>CASE</th>
<th>CONC</th>
<th>&lt;10D</th>
<th>DATE</th>
<th>GROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Descriptive Statistics
- Mean = 22.2000
- SD = 37.1000
- GM = 3.1960
- GSD = 11.920

Compliance Statistics (lognormal)
- X0.95 = 242.0000
- 95%LCL = 17.1000
- 95%UCL = 1.818009
- IncFrac = 0.095
- 95%LCL = 0.004
- 95%UCL = 0.551

Compliance Statistics (non-parametric)

Bayesian Decision Charts
Type of prior decision distribution: Uniform prior

<table>
<thead>
<tr>
<th>Rating</th>
<th>0-7</th>
<th>1-HC</th>
<th>2-WC</th>
<th>3-C</th>
<th>4-PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cutoff (+OEL): 1.0</td>
<td>10.0</td>
<td>50.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Prior
- n = n = n = n = n = n
- Likelihood 0.000 0.107 0.716 0.143 0.035
- Posterior 0.000 0.107 0.716 0.143 0.035
- Cum Likelihood 0.000 0.107 0.826 0.969 1.000
- Cum Posterior 0.000 0.107 0.826 0.969 1.000
What happens when our sample GSD exceeds our GSDmax parameter?
Adjust the “Universe” to account for a larger GSD…

Notice that the Max Likelihood GSD is now in parameter space!
What do we do now?

• What might be going on with Process Operator #2?
• Which tasks might be creating the issues?
• Should we institute a task-based sampling strategy? Which tasks?
• Wildly disparate data result in extreme and unlikely sample GSDs, pushing the decision probabilities toward the higher Ratings.

• Possible solutions:
  – Separate the data and analyze separately.
  – Replace low measurements with higher LODs.
  – Collect more data.
Informative Prior Based On Past Monitoring

• Leveraging Monitoring Data From Similar Operations
Example

• Process equipment being relocated from Brazil to India. The same engineering controls are installed in the new facility in India.

• Lets utilize past sampling data from Process Operator #1 (Brazil) to construct a custom prior for our new Process Operator #1 (India).
Decision Charts

- Process Operator #1 (xylene)
- \( n = 4 \),
- OEL=100 ppm
- \( x = \{13,26,18,12\} \) ppm

- Use the “Likelihood” chart as the new “Custom Prior”
Process Operator #1 (India)

- Custom Prior was constructed with data from Brazil to be leveraged for India
Comments

• The Prior Decision Chart has a greater influence on the Posterior Decision whenever the sample size is small.

• For large sample sizes, say n>10, the Prior has less influence on the Posterior. But for Category 4 it can still be significant!

• Consequently, the accuracy of the Initial Rating is a critical issue whenever the sample size is small.
Impact of Prior on Small & Medium* Size Datasets

Data Sets: **Sampling Data = Category 2 (10-50% of OEL)**

\[ X = \{12, 21\} \]

\[ X = \{12, 21, 14, 11, 18, 9, 24, 26\} \]

* - We will consider 8 data points a medium size dataset for this exercise.
Descriptive Statistics
Mean = 16.5000
SD = 6.3600
GM = 15.9000
GSD = 1.485

Compliance Statistics (lognormal)
X₀.₉₅ = 30.4000
95%LCL = 19.2000
95%UCL = 5.17E0005

Prior & Data Category Match
Prior Cat = 2
Data Cat = 2

Descriptive Statistics
Mean = 16.9000
SD = 6.3300
GM = 15.8000
GSD = 1.475

Compliance Statistics (lognormal)
X₀.₉₅ = 30.0000
95%LCL = 23.0000
95%UCL = 54.6000
Descriptive Statistics
Mean    = 16.5000
SD      = 6.3600
GM      = 15.9000
GSD     = 1.485
Compliance Statistics (lognormal)
X0.95   = 30.4000
95%LCL  = 19.2000
95%UCL  = 5.17E0005

Prior & Data Category Mismatch!
Prior = Cat 4
Data = Cat 2

Note how “n” impacts Final Decision
n=8
n=2
Impact of Prior on Small & Medium* Size Datasets

Data Sets:

Sampling Data = Category 4 (>100% of OEL)

\[ X = \{65, 29\} \]

\[ X = \{65, 29, 48, 108, 42, 33, 16, 57\} \]

* - We will consider 8 data points a medium size dataset for this exercise.
Descriptive Statistics
Mean = 47.0000
SD = 25.5000
GM = 43.4000
GSD = 1.770

Compliance Statistics (lognormal)
Mean = 49.8000
X0.95 = 111.0000
SD = 28.3000
95%LCL = 75.1000
GM = 43.3000
95%UCL = 270.0000
GSD = 1.777

Prior & Data Category Mismatch!
Descriptive Statistics
Mean = 49.8000
SD = 28.3000
GM = 43.3000
GSD = 1.777

Compliance Statistics (lognormal)
X0.95 = 111.0000
95%LCL = 75.1000
GM = 43.3000
95%UCL = 270.0000
GSD = 1.777

Note Very Low % in Cat 4...
THIS CREATES A BAD RESULT!!!

Descriptive Statistics
Mean = 47.0000
SD = 25.5000
GM = 43.4000
GSD = 1.770

Compliance Statistics (lognormal)
X0.95 = 111.0000
95%LCL = 56.9000
95%UCL = 1.4E0008
Impact of Mismatched Prior

We get penalized for a mismatching Prior & Sampling Data
Warnings on Creating Priors to be leveraged across SEGs

• An incorrect prior can drive the wrong decision in some circumstances
  – Careful when putting a very low % in any one category of a prior

• Important to create a process for “validating” priors using sampling data from same SEG
  – Minimum # of Samples
  – Universe GSD boundaries / Max sample GSDs
  – Rules on task differences
  – Rules on engineering controls
Conclusions
Industrial Hygienists Are Bayesian Thinkers!
What data should we use for the Likelihood Decision Distribution?

- Current data (<2 years old)
- Personal exposure data is preferred
- Same equipment & task
When can we use an informative Prior Decision Distribution?

- Informative Priors can be used whenever we are confident in our Professional Judgment.
- Professional Judgment can be based upon…
  - Past experience with this or similar processes or tasks
  - Analysis of fairly recent data
  - Physical/chemical modeling

- Use a Uniform Prior if in doubt about the accuracy of Professional Judgment.