Models vs. Measurement: Are we on the right Track

Noah Seixas

IOHA/BOHS
London
April, 2015

Search: “exposure models”
result: exposure.inc

What is measurement data?

- Information
  - Random observations, notes and measurement values
  - Have little interpretability, but when organized becomes:
- Data
  - Organized datasets support:
- Analysis
  - Demonstrates patterns, associations, relationships
  - Allows for:
- Interpretation
  - Ability to answer a question or tell a story

All the above, determined by the story we want to tell.

Problems with data

- Data are expensive to collect
  - Requires technical expertise
  - Requires specialized equipment
- Data have high variability
  - Therefore, uncertain
  - Requires a large sample size (expense)
- They represent the present
  - Can’t measure the past
  - Future will certainly be different

Solution: Use a Model

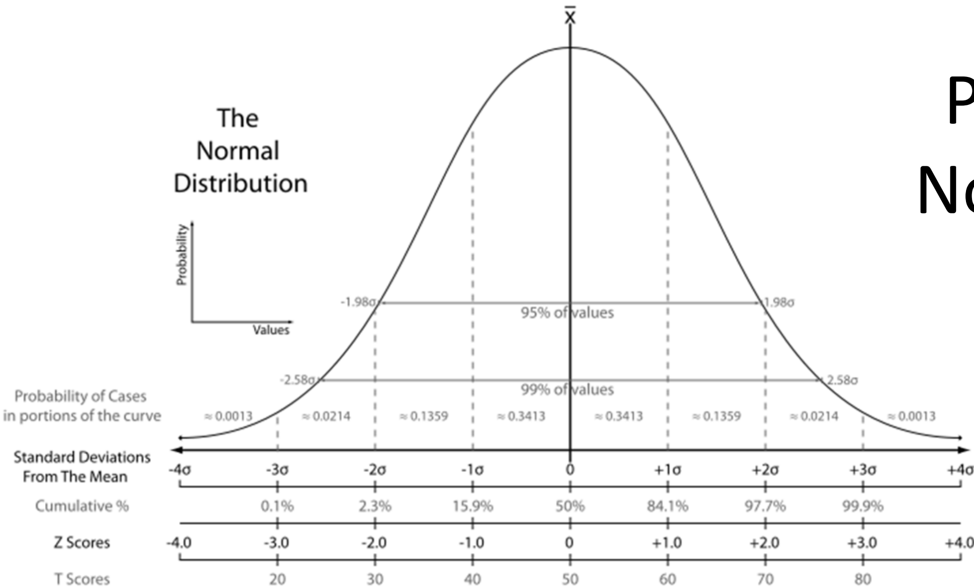
What is a model?

- A physical, mathematical or mental construct allowing us to describe our expectations
  - Estimation of conditions or relationships
  - Prediction of unobserved conditions

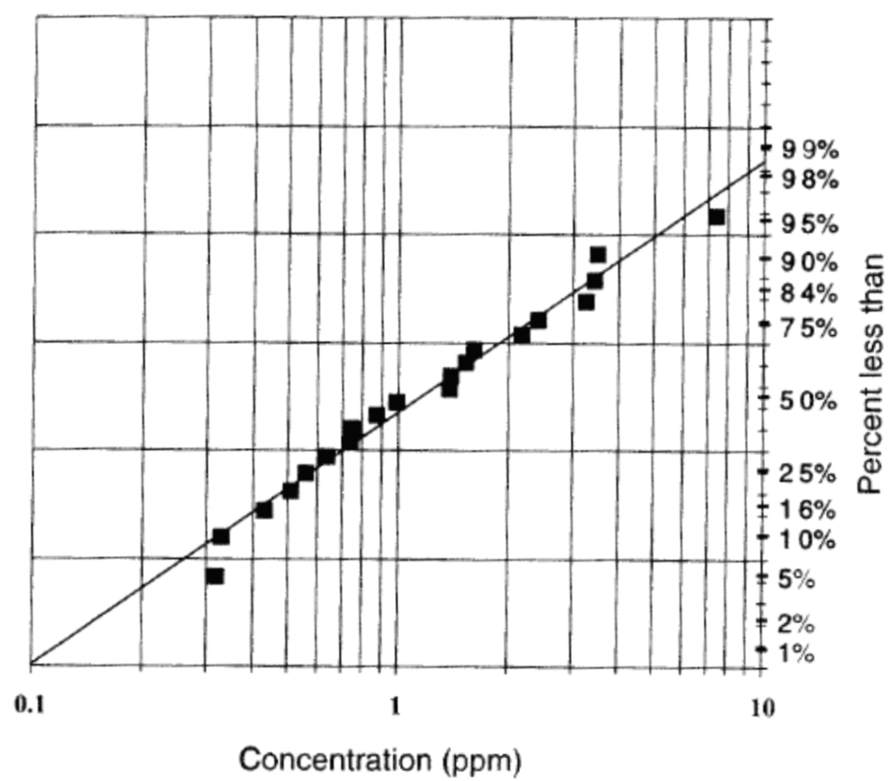
The lognormal model for exposure distributions

- Supports estimation
  - Estimation of parameters with confidence intervals
  - Know what we mean by GM, GSD, 90^{th} percentile
- Supports prediction
  - Given our model, what is the probability of a samples exceeding the OEL, etc.

Parameters of the Normal Distribution



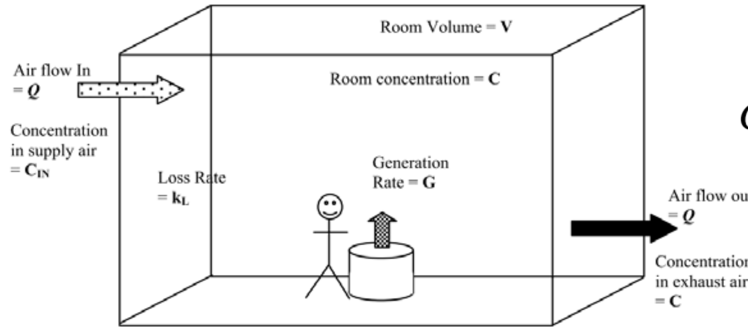
Log-probability plot for estimation or prediction of LN parameters (from Ramachandran, 2005)



Types of Exposure Models

- Deterministic
  - E.g., well mixed box

(From: Ramachandran, 2005)



$$C_{\text{steady state}} = \frac{G + C_{\text{IN}}\,Q}{Q + k_{\text{L}}\,V}$$

- "Mechanistic"
  - E.g., ART

(From Fransman, et al, AOH, 2011)

$$C_t = \left[\frac{1}{t_{total}} \sum_{tasks} \left\{ t \cdot (C_{nf} + C_{ff} + Su) \right\} + t_{non-exposure} \cdot 0 \right] \cdot RPE$$

$$C_{nf} = (E_{nf} \cdot H_{nf} \cdot LC_{nf1} \cdot LC_{nf2}) \cdot D_{nf}$$

$$C_{ff} = (E_{ff} \cdot H_{ff} \cdot LC_{ff1} \cdot LC_{ff2} \cdot Seg_{ff}) \cdot D_{ff} \cdot Sep$$

$$Su = Su_{factor} \cdot (E_{ff} \cdot H_{ff} \cdot LC_{ff1} \cdot LC_{ff2} \cdot Seg_{ff} \cdot D_{nf} \cdot Sep_{ff})$$

- Stochastic
  - E.g., mixed effects

(From Friesen, et al, AOH, 2015)

$$\begin{aligned} \text{Ln}(Y_{\text{adj}})_{jifd} = \beta_0 &+ \sum \beta T + \sum \beta I_{\text{job}} \\ &+ \sum \beta (I_{\text{job}} \times I_{\text{ind}}) + bJ_j \\ &+ bInd(J)_{ji} + bO_{jif} + \varepsilon_{jifd}. \end{aligned}$$

Data in models

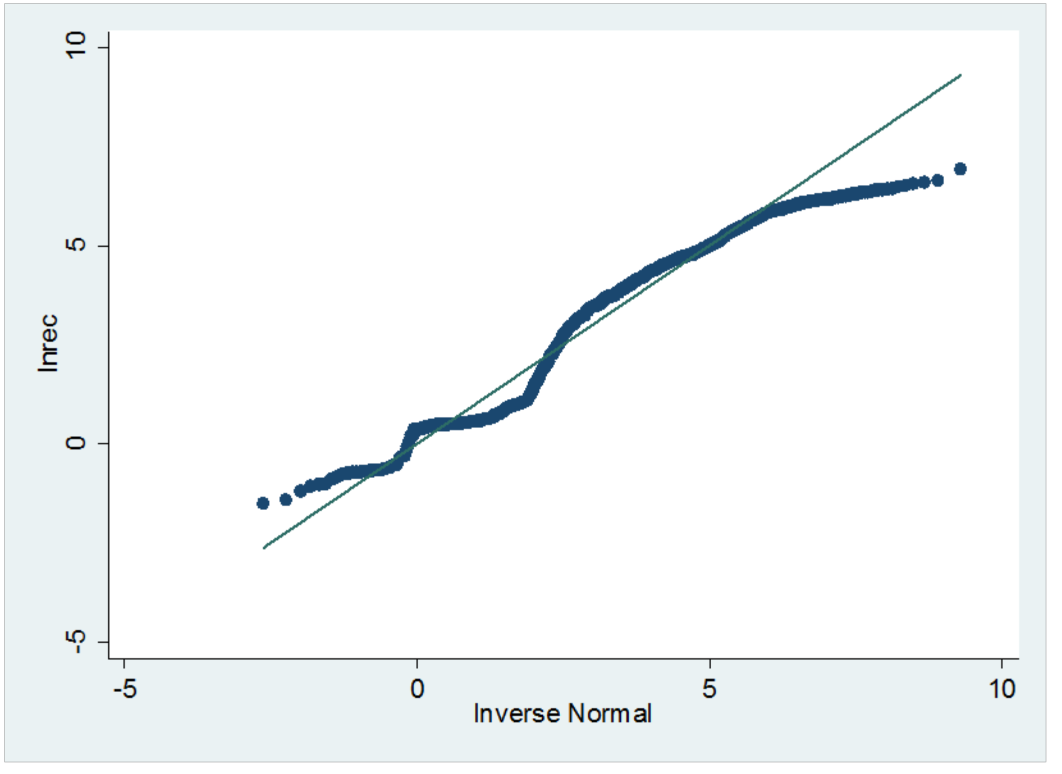
- Deterministic
  - Based on chemical and physical principles
  - Exposure data not required
- Mechanistic
  - Exposure data used to develop factors
  - Also used to 'calibrate' model outputs to specific scenarios
- Stochastic
  - Based entirely on observed exposure data
  - And observed determinants

All types require exposure data for validation

Use of models

- Despite very different levels of data input all models function in a similar fashion:
  - Can estimate or predict an expected value
  - Accuracy depends on
    - ‘Correctness’ of model structure
    - Quantity and quality of data used in development
  - In other words: the real world may not be ‘modeled’ correctly

Real world may deviate from our model, e.g., LN



Respirable Elemental Carbon in Underground Mines
(n=1256)
from NIOSH/NCI DEMS Study

Advantages of models/measurement Depends on Purpose

	Management		**Epidemiology**		**Risk Assessment**
	Control Banding	Compliance	Retrospective	Prospective	Population Distribution
Deterministic					
Mechanistic					
Stochastic					
Data					
Excellent					
OK					
Caution					
Poor					

What’s a hygienist to do?

- Models best for expected conditions
  - Or where data are missing
- Measurement data required to
  - Develop and validate models
  - Identify deviations from expectations
  - Consider particular situations
    - e.g., OEL compliance
- Combined use of model and data
  - Exploit the strengths of both approaches
    - Robustness of model, particularities of a given condition
  - Bayesian approach
    - Model: prior
    - Data: update
    - Posterior: combination of two

Conclusions

- Measurement data are required
  - Foundation of OH science
- Models can draw significant strength from data
  - Not in absence of data
  - Support our understanding of data
- Measurement data always required for validation
  - Otherwise, we don't know if the model is 'correct'
- Trend toward fewer data will undermine our models, and our understanding

Discussion

- Main question:
  - How can we best use models while continuing to stimulate quantitative exposure assessment?

- To what degree can deterministic, mechanistic or stochastic models replace exposure measurements?
  - When?
  - Where?
  - Which?
- How can we assure that models are not mis-used?
- Should all occupational hygienists be comfortable using and interpreting models?
  - Deterministic
  - Mechanistic
  - Stochastic