Relating Data and Models to Characterize Parameter and Prediction Uncertainty

---

R. Woodrow Setzer

US EPA

ORD/National Center for Computational Toxicology

Research Triangle Park, NC

919-541-0128

setzer.woodrow@epa.gov

Outline

- Preface
- Differentiating Variability from Uncertainty, and consequences of both.
- Uncertainty about parameter values
  - Uncertainty of prior information
  - Uncertainty after estimating using PK data
- Evaluating Model & Dose-Metric Uncertainty

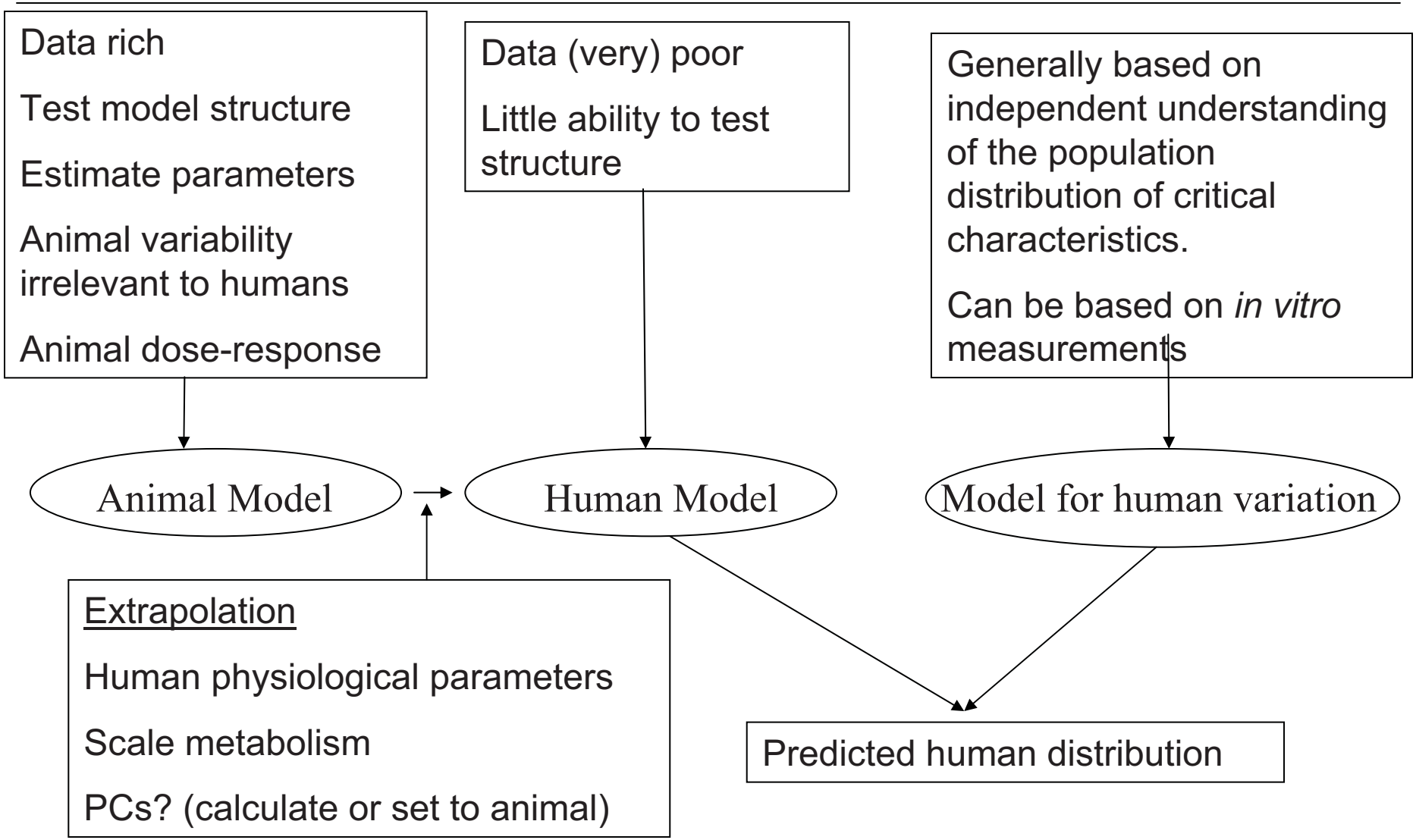
Abbreviations

- CV – coefficient of variation = standard deviation / mean
- MC – Monte Carlo
- MCMC – Markov Chain Monte Carlo
- ML – maximum likelihood
- PC – partition coefficient
- TCE – trichloroethylene

Preface

- Level of treatment: overview of what needs to be considered, not a ‘how to’
- Methods need
  - Objectivity: decisions are based on pre-specified criteria
  - Transparency: decisions and computations are auditable
  - Replicability: results can be reasonably replicated by a reasonably competent practitioner
  - Verisimilitude: we can be reasonably confident that answers are at least approximately “right”
- Presumption:
  animal model → human model → human predictions

Path to Human Dosimetry Predictions



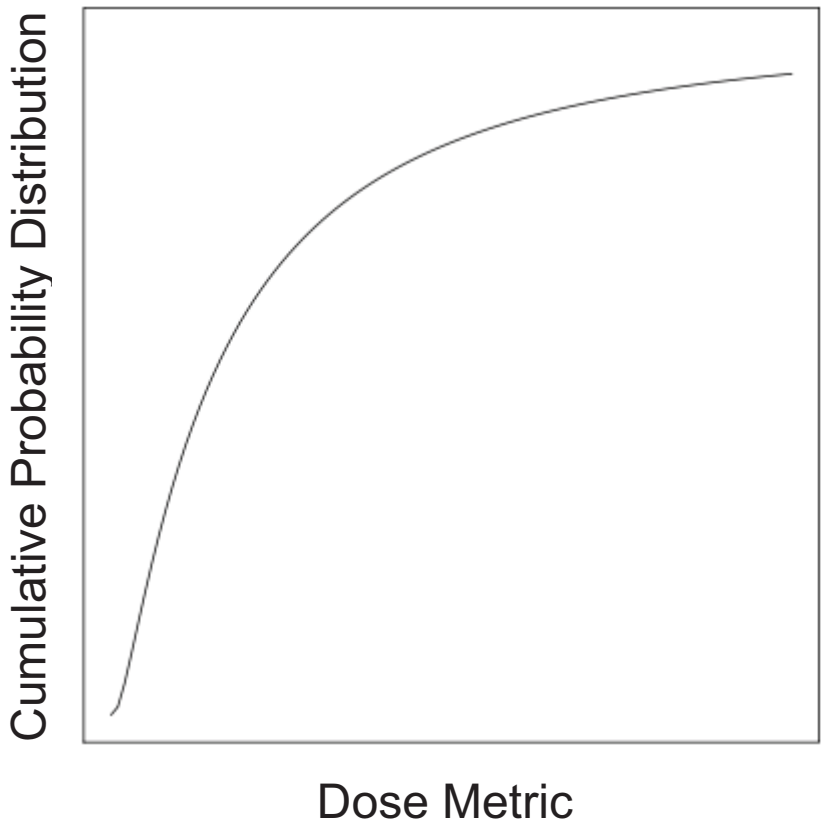
Variability and Uncertainty

Variability

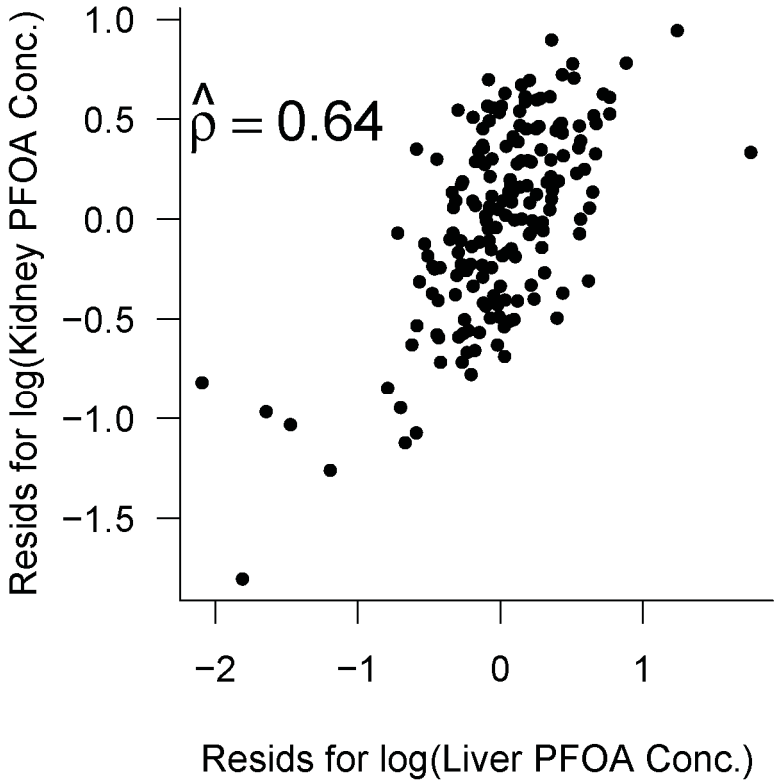
- *Variability* is a characteristic of a *population* of observations:
  - Height, weight, IQ
  - Clearance, liver blood flow
  - Receptor or enzyme concentration
  - Measurements
- Individuals in the population differ with respect to the characteristic in question
- We typically use the language of probability distributions and random variables to model variability (*e.g.*, variance)

Practical Consequence: The Population Distribution of the Dose-Metric

- Variation among exposure scenarios
- Variation in characteristics of target population (pharmacokinetics):
  - age
  - sex
  - body weight
  - activity level
  - metabolism
  - cardiac output
  - other biochemical (binding, etc.)



Practical Consequence: Interindividual Variation in Experimental PK Studies



Tissue concentrations within individuals may be correlated, which may affect parameter estimates and the way we should evaluate their uncertainties.

Residuals from regression of log(liver) and log(kidney) on time, block, gender, and interactions, 10 mg/kg PFOA by gavage in mice. Lou et al, *in press*

The Human Variability Model

- What varies?
  - Physiology and anatomy
    - Predictors
      - Age
      - Gender
      - Body weight
      - Disease state
      - …
  - Biochemical variability:
    - Enzyme polymorphisms (but not all!)
    - more complex variation of enzyme activity, like continuous variation due to induction, health status, etc.
    - Transporters
    - Binding proteins
- (Joint) Distributions?
- How do we know these values?

…How Do We Know these Values?

- Samples of convenience
- Measured in 9 healthy young male volunteers
- Collections of microsomes or other tissues
- Stratified random sample of the relevant target population with defined probability of inclusion (hah!)

- Uncertainty comes from:
  - Estimation from a sample
  - Relevance of the sample to the target population
  - Relevance of the estimate to the target parameter

Uncertainty

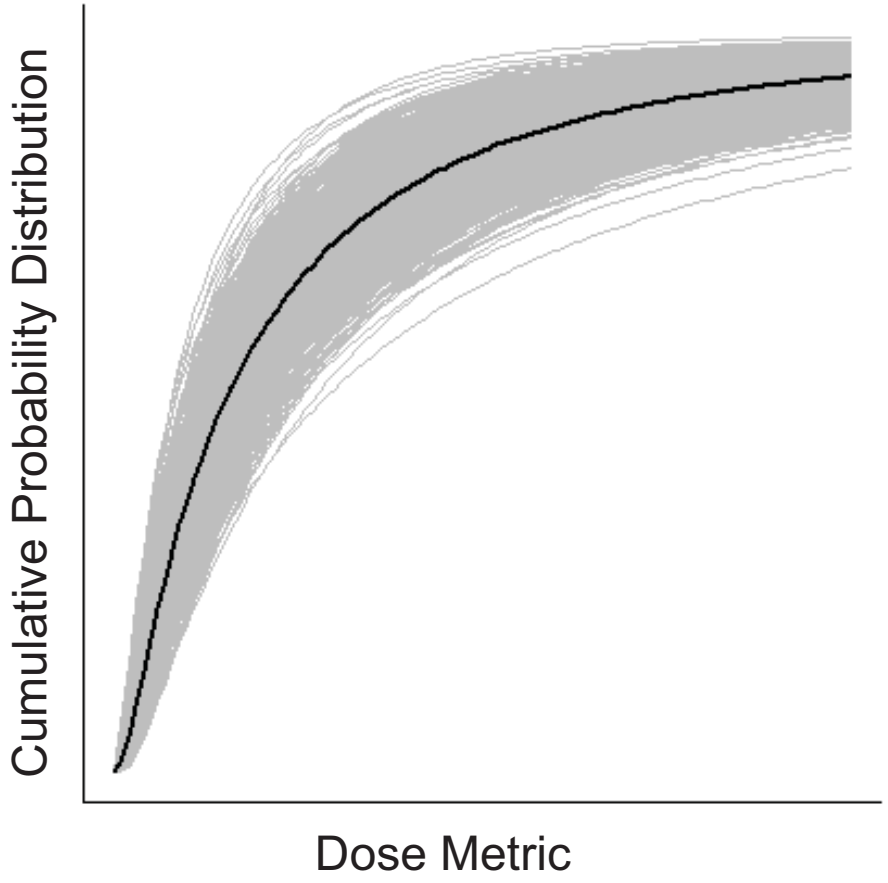
- *Uncertainty* is a characteristic of a single observation:
  - Parameter estimates (model parameters, sample means)
  - Predictions from models with uncertain components
  - Individual samples from a population with variance
    - E.g. cardiac output for an individual (rat or human) characterized by the population mean.
- Usually can reduce uncertainty with more data (at least when speaking of parameter uncertainty).
- Can be *uncertain* about variability: *e.g.*, suppose we are uncertain about the precise value of the variance of a characteristic in a population.
- Usually use the language of probability for uncertainty, but can also be more qualitative: intervals, bounds.
- 'Uncertainty' also applies to more qualitative entities: model structure, mode of action, etc.

Expressing Uncertainty about Variability

Don’t know:

- variability of exposure details
- pharmacokinetic parameters
- population distribution of varying characteristics
- True pharmacokinetics
- …

So, characterize the uncertainty in the population distribution of the dose-metric



Kinds of Uncertainty

- Uncertainty about model parameters
  - Easiest to quantify, but often the smallest contribution.
- Uncertainty about mode of action
  - What is the appropriate dose metric?
- Uncertainty about the biology
  - What is the appropriate model structure?

(After Clewell, 2006)

Parameter Uncertainty

Parameter Uncertainty is Tied to How the Parameter is Estimated

- "Assumed", based on species-specific conventions
  - Physiological parameters
- In vitro →in vivo
  - PCs
  - Metabolic parameters
  - Protein binding
  - Dermal absorption
- Computed
  - PCs
  - Binding(?)
- Extrapolated
  - Body weight scaling for metabolic parameters
- Estimated from PK data using model
- Directly measured

Uncertainty about Assumed Parameters

- The true value for an individual is unknown, but there are estimates of population “central tendency”, and the population variance is presumed “small”.
- The uncertainty is the population variance + uncertainty about the population mean
- Correlations – dependencies?
- Physiologically realistic bounds

Example

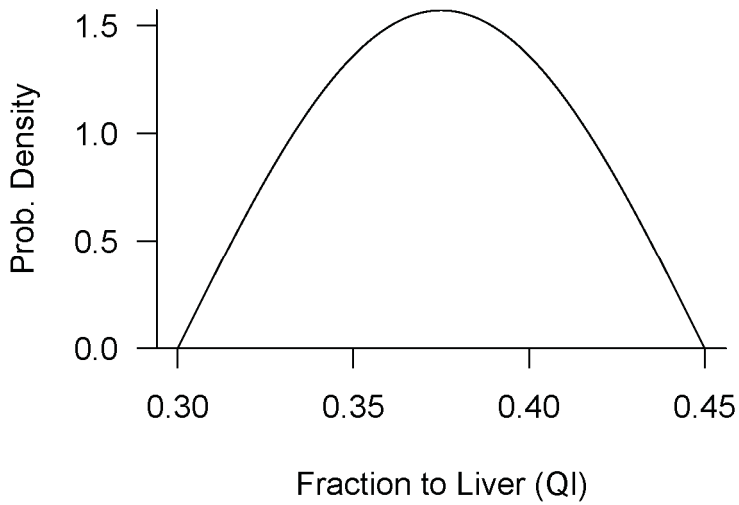
in analysis of model for TCE:

Tissue blood flow is fraction of total (constrain so fractions always add to 1)

Q_l, Q_f, Q_p ~ Bounded cosine distribution

$Q_r = 1 - (Q_l + Q_f + Q_p)$

Bounds are median plus or minus 20% (experimenters' intuition)

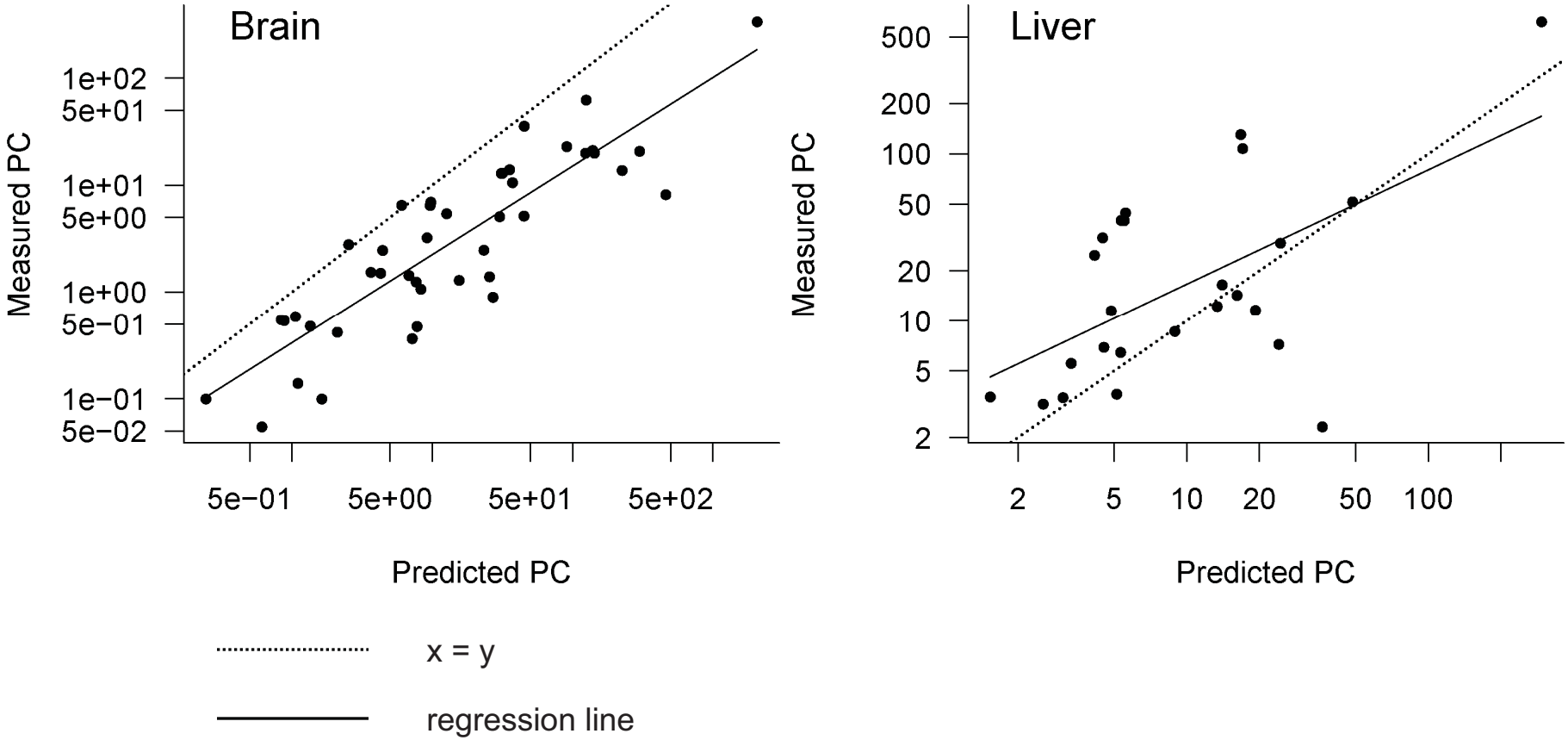


(from Bois, et al. 1990)

Computed

- For example, computed PCs as in, Pelekis, et al (1995), Poulin and Thiel (2000), Rogers and Rowland (2006), and Schmidt (2008)
- Uncertainty: compute mean squared error for prediction of (log) PC values, or, variance around regression of measured on predicted PC values.

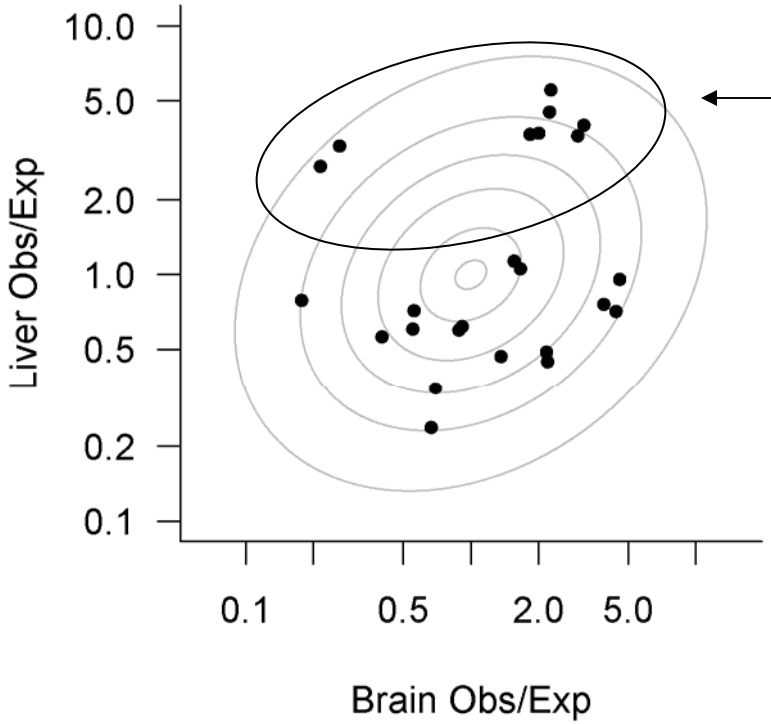
Comparison to Observed Values

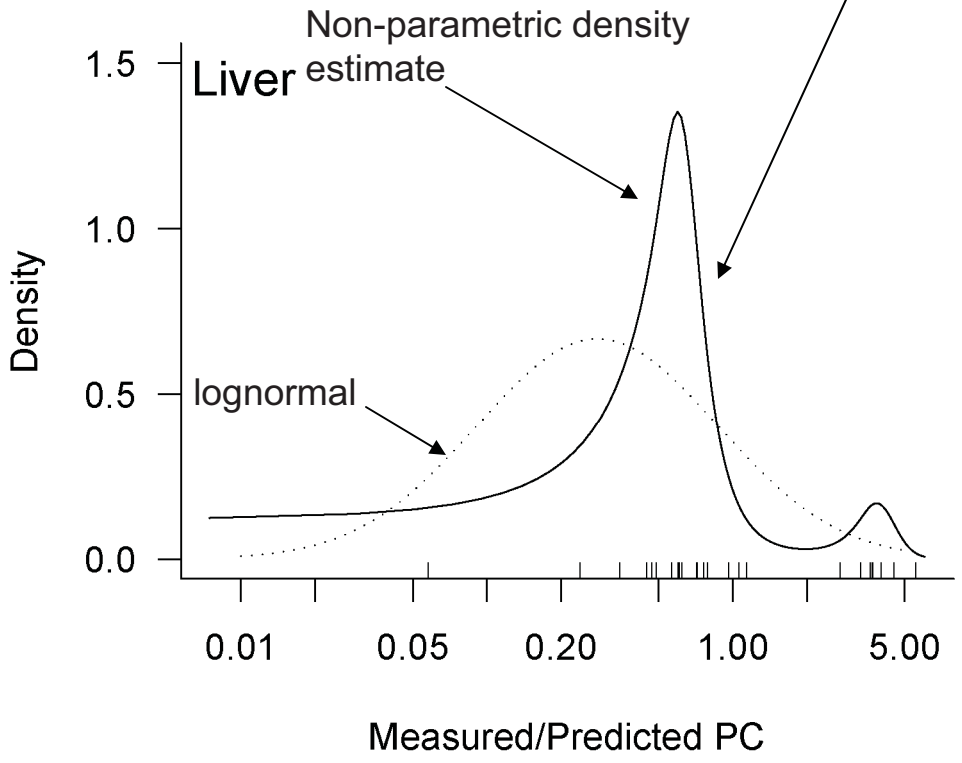


Distribution of Errors: Uncertainty Distribution for PCs

More Mature Distribution

- Joint distribution of errors?
- Shape of error distribution?
- Other predictors?





Estimated from PK Data Using a PBPK Model

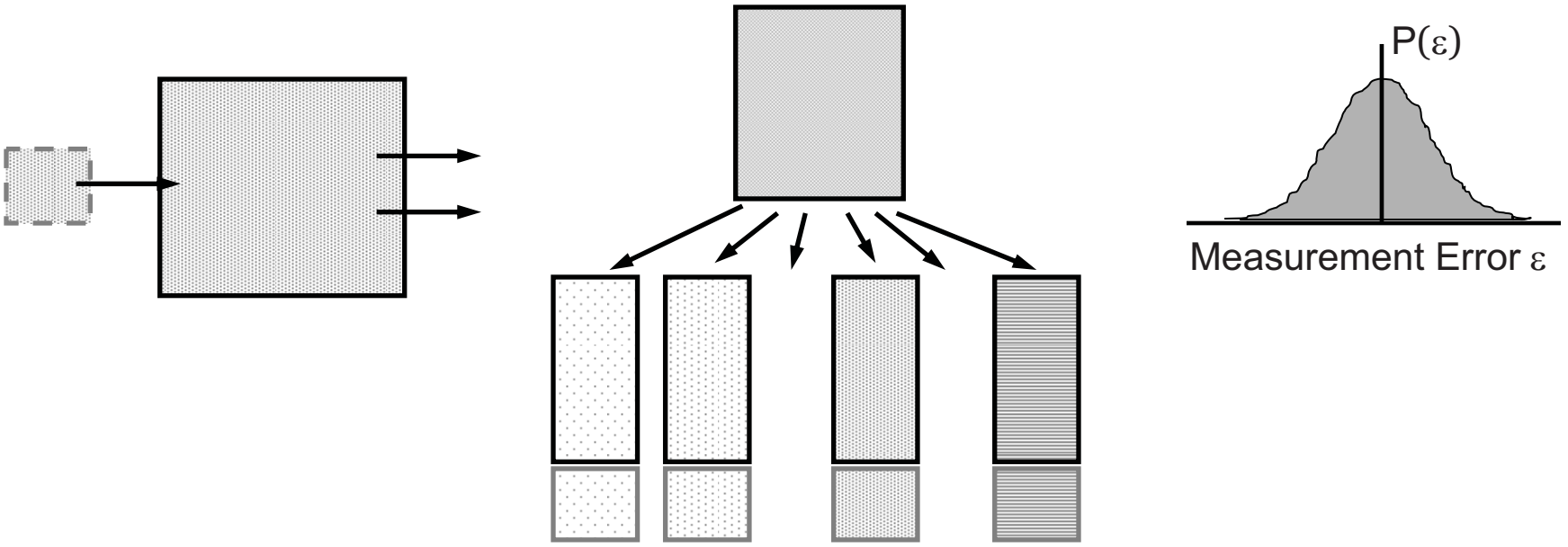
- Several estimation methods are commonly used:
  - Visual fit: no quantitative or objective measure of uncertainty
  - (Weighted) least squares: standard errors, etc. only valid when observations are all independent (rare with PK data).
  - ***Maximum likelihood***: standard errors are valid if the model is (close to) correct, but assume non-estimated parameters are known exactly.
  - ***Bayesian methods***: Can accommodate partial knowledge about all parameters.

Models for full characterization of variability and uncertainty

PK and/or PD Model

Hierarchical Statistical Model

Measurement Model



Estimate all parameters simultaneously.

The Likelihood Function

- Basis for *all* inference about models.
- Combines model and data in a single expression.
- Depends on
  - Experimental design
  - Nature of variation
- Example (lognormal error model):

$$L(\mathbf{Parms} \mid \mathbf{t}, \mathbf{Inputs}, \mathbf{y}) = \sigma^{-n} \prod_{i=1}^{n} \frac{1}{y_i} \mathrm{e}^{\frac{-(\ln(y_i)-\ln(F(t_i, \mathbf{Inputs}_i, \mathbf{Parms}))^2}{2\sigma^2}}$$

Maximum Likelihood

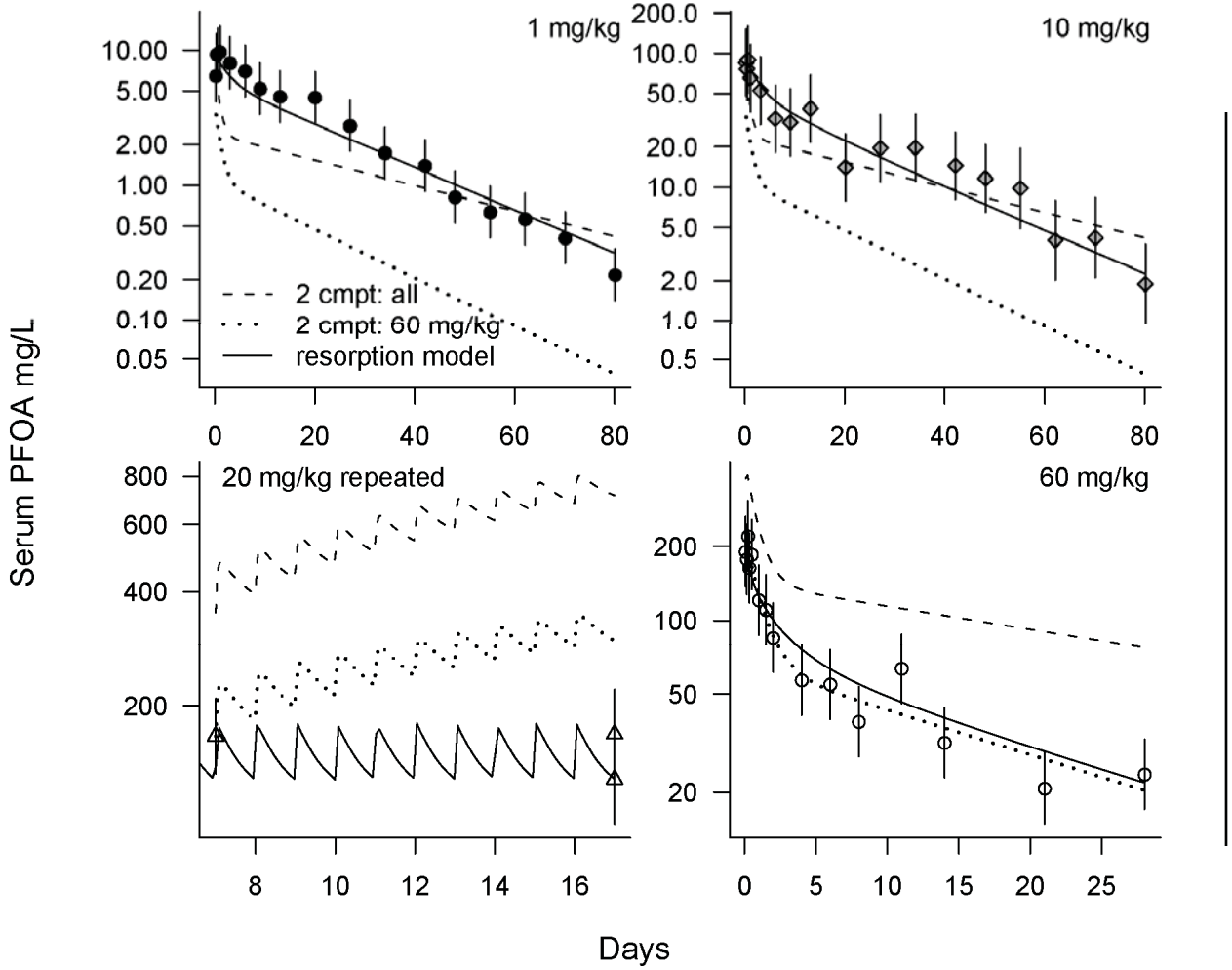
- Construct a likelihood function for the data, based on the statistical model.
- Use iterative numerical methods to find the parameter values where the likelihood is a maximum. ML requires optimization.
- Estimates are approximately normally distributed (transformation, like log, can improve the approximation).
- Covariances among the estimates can be estimated.
- Models of population variability are numerically challenging.
- Models with many estimated parameters are difficult.
- Parameters are either estimated, or fixed.

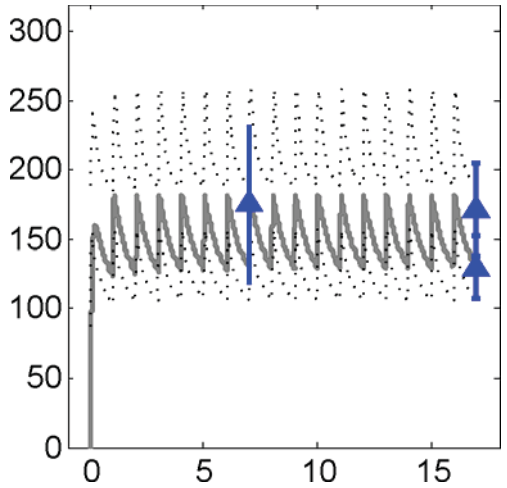
Example: ML Estimation for a Biologically Motivated PK Model

- Point was to compare 1 & 2 compartment models and 2 compartment model with saturable resorption.
- Individual data (no repeated measures), single doses of 1, 10, 60 mg/kg, 17 day repeated dose study at 20 mg/kg in mice.
- Serum concentrations
- Assume lognormal errors.

Lou et al. (2008)

PFOA Fits



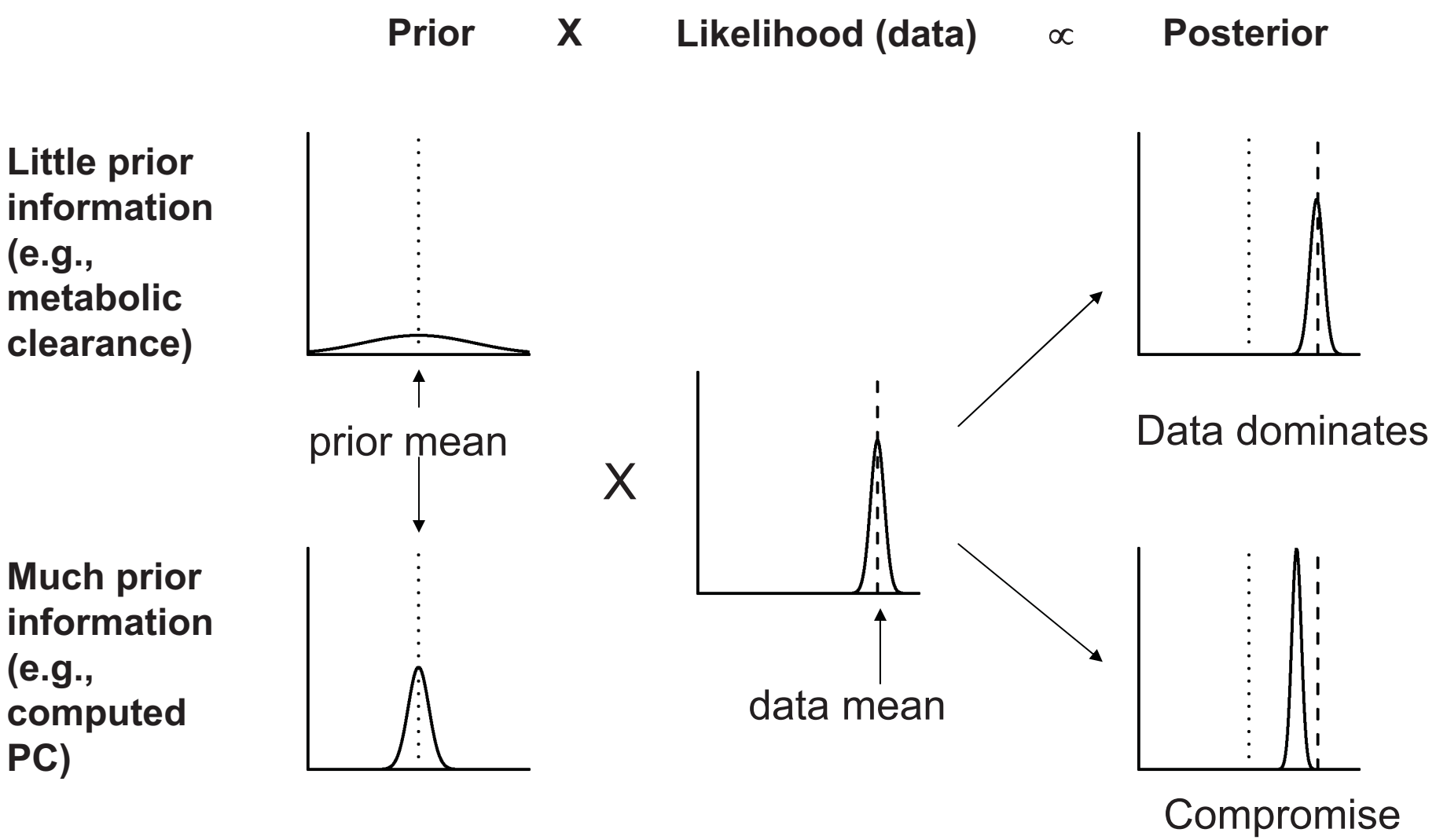


Dashed lines indicate uncertainty in model predictions based on ML fit.

Bayesian Methods Combine Prior Information and Likelihood

- Baysian analysis allows information about uncertainly known values (e.g., partition coefficients, metabolic parameters) to be included as uncertain (in ML, we either estimate parameters, or treat them as known perfectly).
- Relatively simple treatment of population models (including designs with repeated observations of the same animals).
- The goal of Baysian analysis is to calculate the *posterior distribution* for all the parameters
  - Not an optimization problem
  - The posterior distribution captures the uncertainty about parameter estimates.
- Generally must estimate the posterior by taking random samples from it using MCMC

Role of Prior in Bayesian Estimation

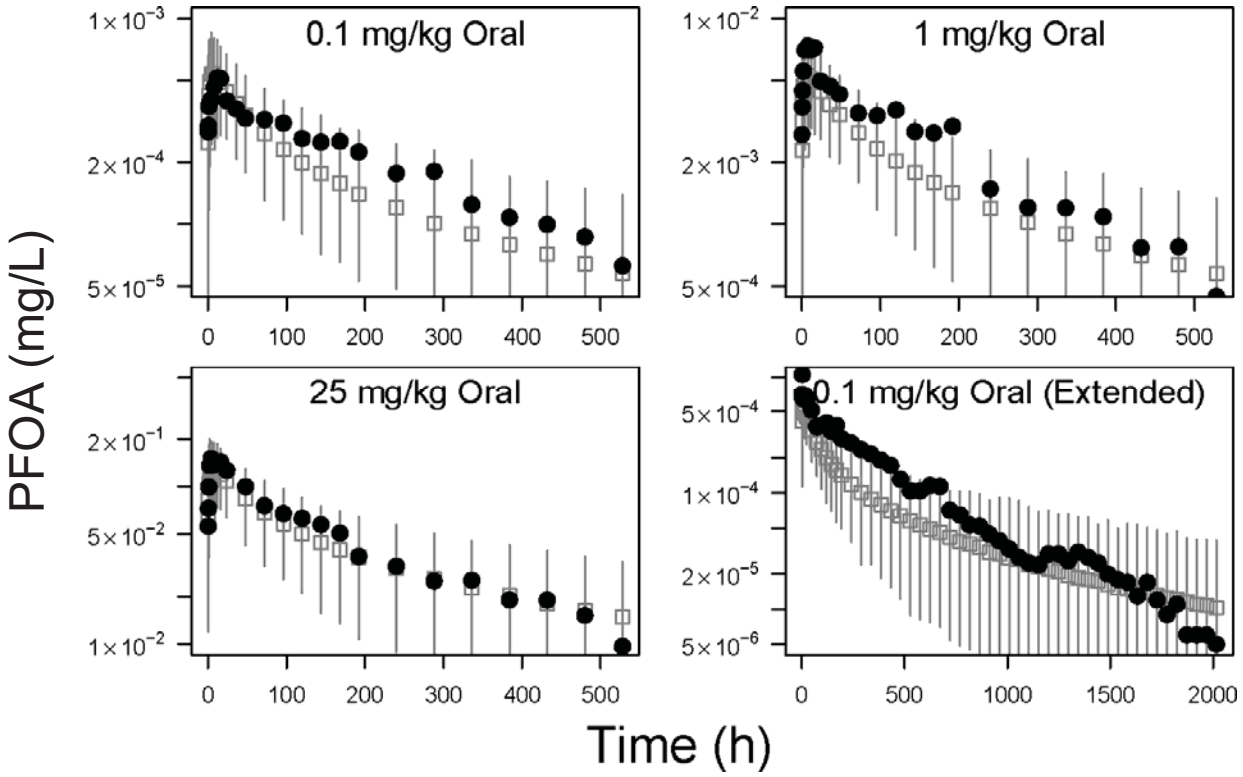


More PFOA

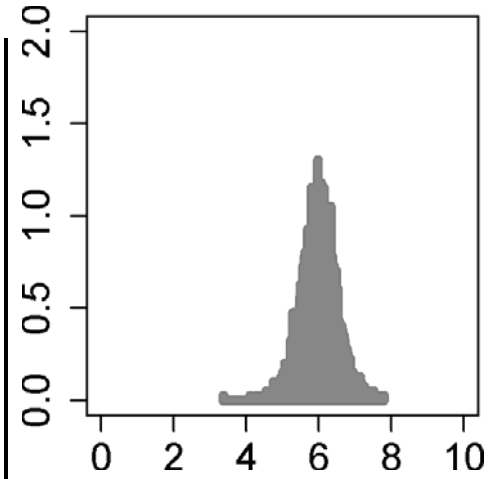
- Multiple studies with repeated observations of PFOA serum levels (rats) after gavage dosing.
- Additional studies, with multiple dose levels, of urine and feces concentrations.
- Two-compartment classical PK models fit; only barely-informative priors.

Wambaugh et al. (2008)

PFOA Fits



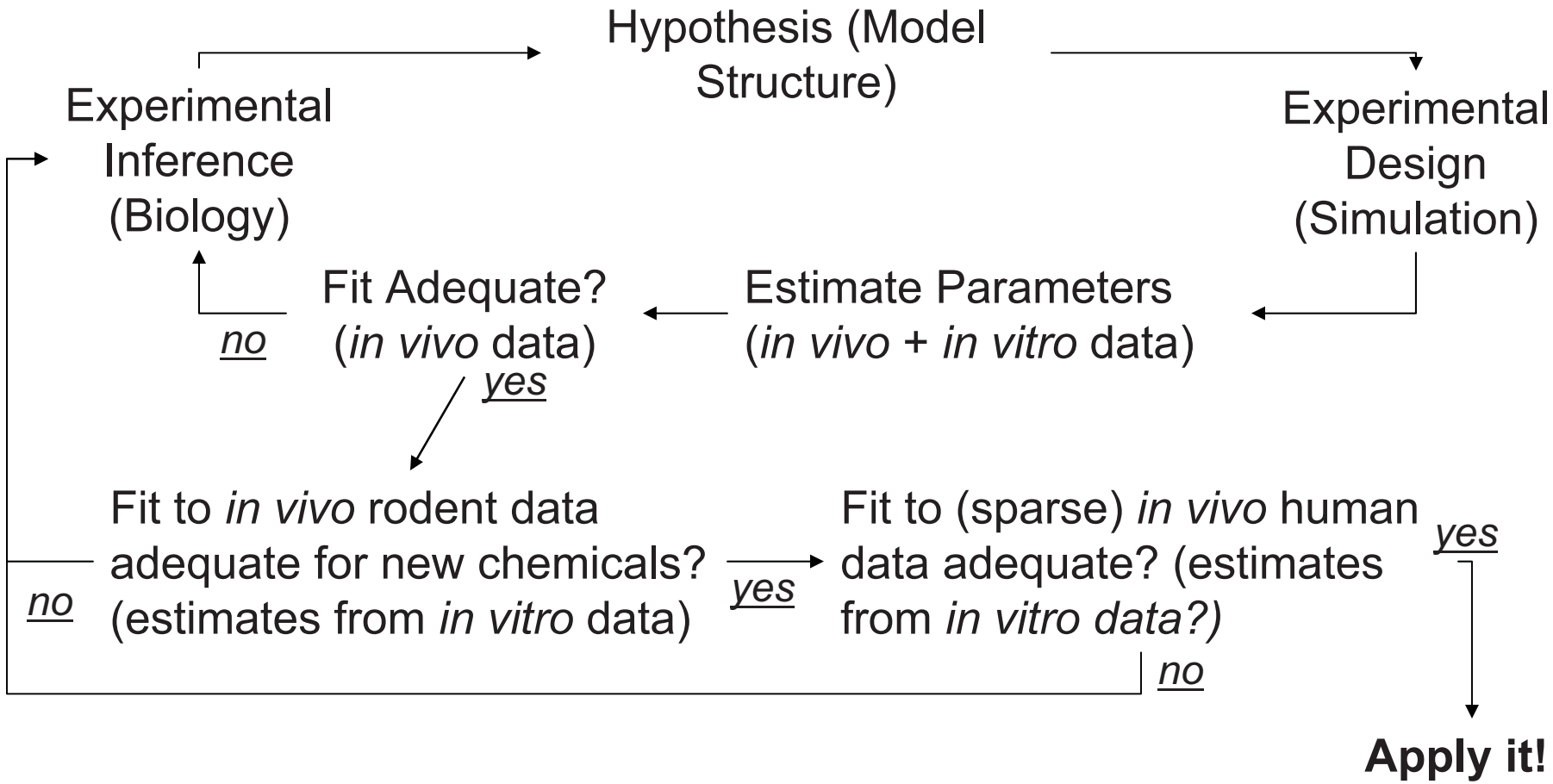
Square symbols and whiskers are means and 95% confidence intervals from model.



Posterior distribution for AUC/dose. From 2 cmpt model, females

Evaluating Model Uncertainty

Roles for Statistical Methods in PBPK Model Development



Model Evaluation

- Plots of scaled residuals (by compartment):
  - Distribution: correct error model?
  - relative to time, dose: deterministic model about right?
  - predicted concentration: right scale for error?

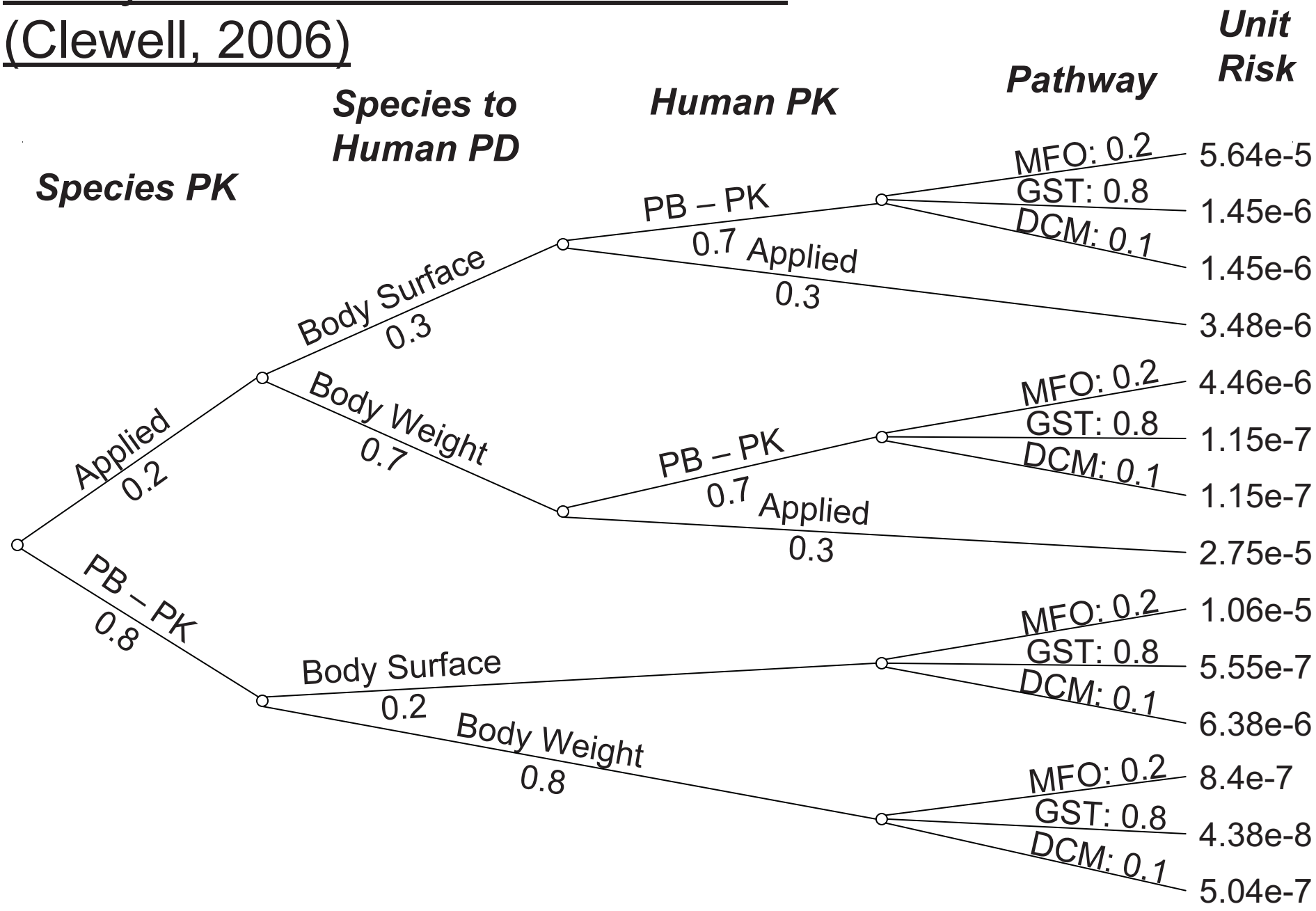
- How different are posteriors from informative priors:
  - *e.g.*, did we have to change a partition coefficient dramatically to fit the data?
  - If very similar, no information in data about parameter

- Posterior correlations between/among parameters
  - Are we estimating individual parameters, or, for example, ratios, as in V_{max} / K_m?

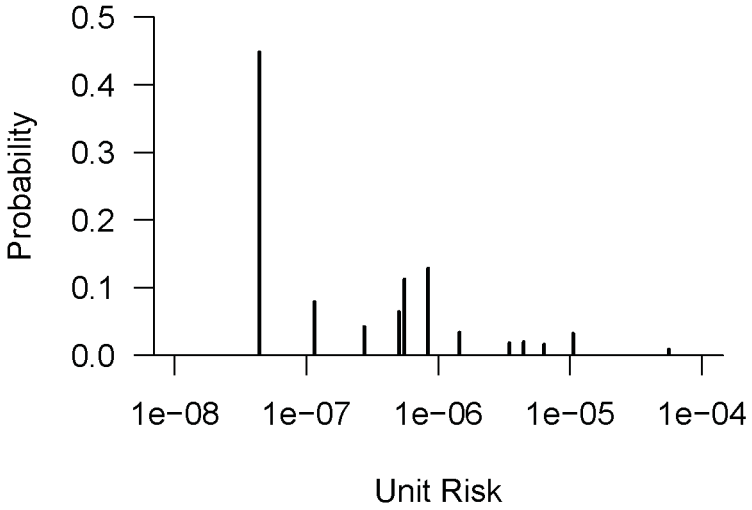
What's the Right Model? What's the Right Dose Metric?

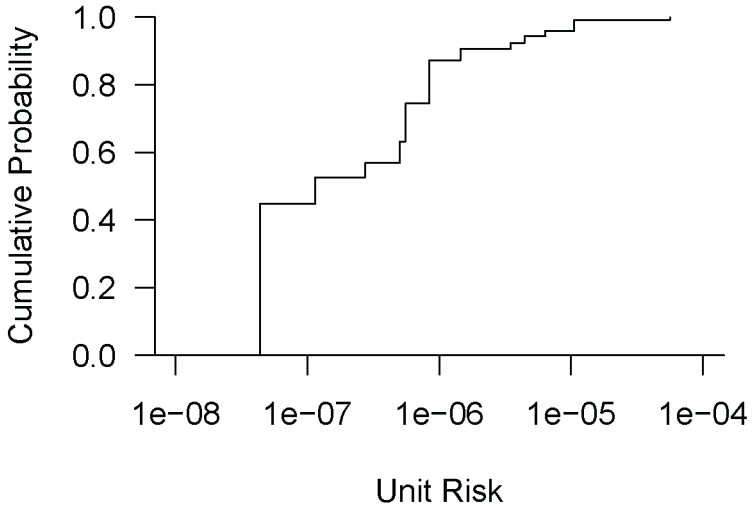
- Model uncertainties:
  - Where is the metabolism?
  - Transporters (presumably saturable) or partitioning?
  - Diffusion or perfusion limited compartments?
  - Complex fat?
  - Inducible metabolism?
  - …
- Can be constructed to give a tree of models
  - Probabilities on branches from expert elicitation, fit to data
- Dose metric (mode of action)
  - Again, can construct a decision tree, with branch probabilities from expert elicitation.

Methylene Chloride Decision Tree (Clewell, 2006)

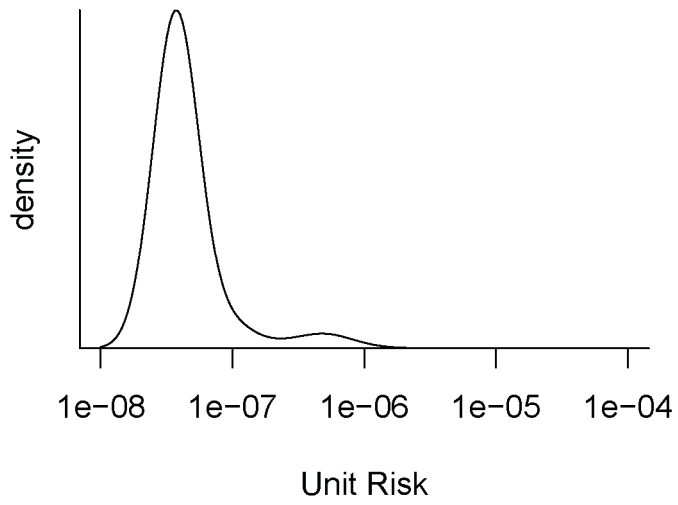


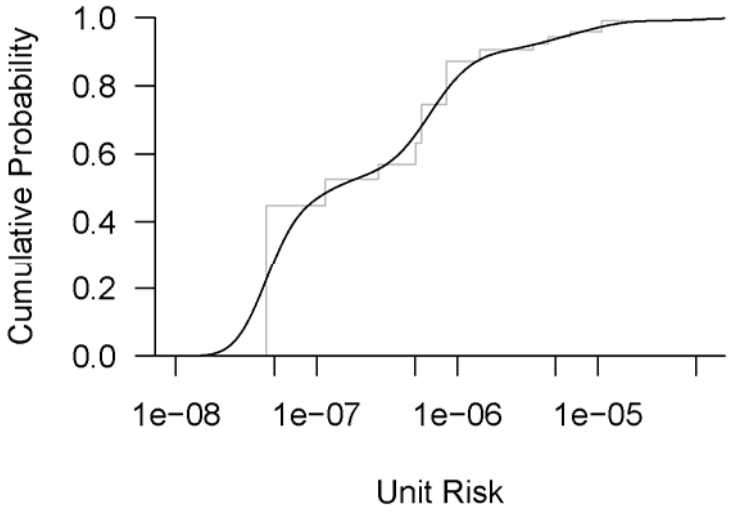
Probability Distribution of Unit Risk Implied by Methylene Chloride Decision Tree:





Usually, estimates of unit risk are uncertain.
If CV for uncertainty distribution is 40%, you get:





Conclusions

- Variability and uncertainty are related but distinct concepts
  - variability deals with populations
  - uncertainty deals with single observations, concepts, parameter estimates.
- A target quantity of a risk assessment using PBPK modeling is often the distribution (“variability”) of a measure of dose, which we estimate (“uncertainty”)
- “Uncertainty” encompasses model uncertainty, uncertainty about the relevant dose metric, as well as uncertainty about model parameters.