Determining an Inflection Point from External-Internal Dose Data

PHILIP VILLANUEVA

HEALTH EFFECTS DIVISION

OFFICE OF PESTICIDE PROGRAMS

Objective

Address problem formulation statement

- There is no guidance on how to analyze data on external-internal dose levels to determine at which measured or statistically-determined external dose levels, the internal doses are significantly non-proportional to external doses

Propose statistical analysis for determining when the relationship between internal doses and external doses significantly depart from proportional

- Piecewise regression with appropriate statistical test and confidence intervals are worth considering

Assumption & Conceptual Model

Relationship between internal doses and external doses is *approximately* linear and proportional at low doses

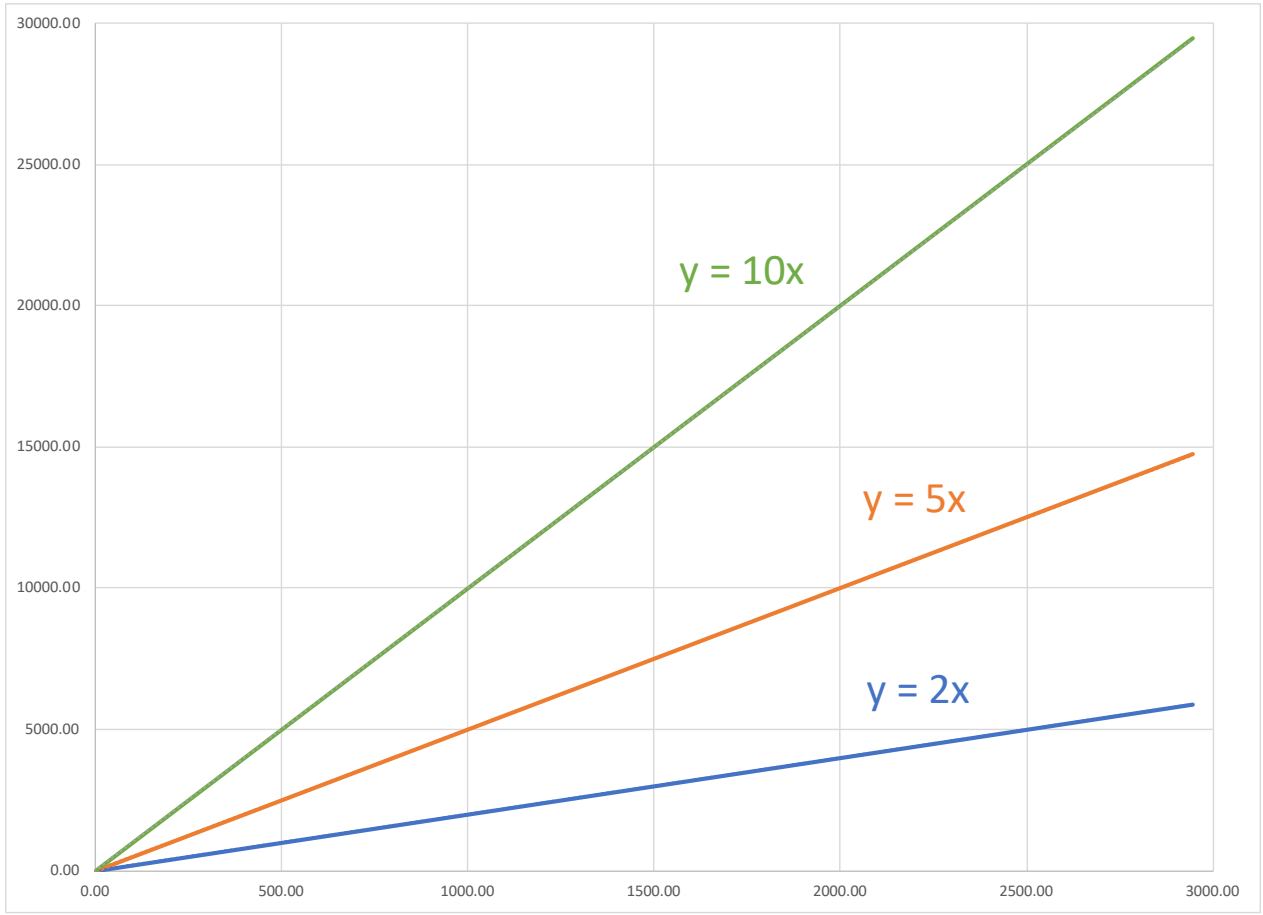
Proportional relationship on linear scale

- $y = \alpha + \beta x$
- Intercept term (α) expected to be zero
- Slope term (β) expected to be proportionality factor

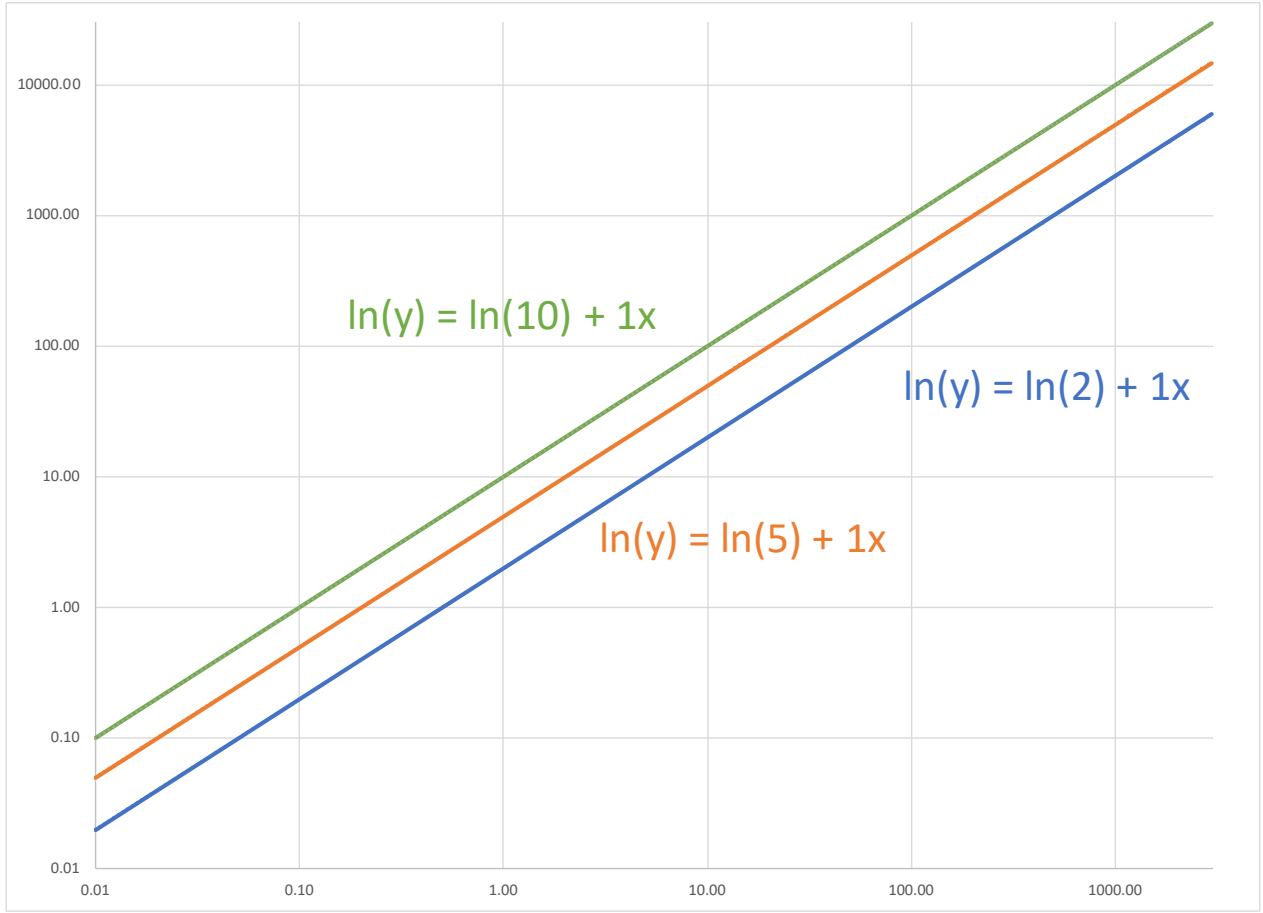
Proportional relationship on log-log scale

- $\ln(y) = \ln(\beta x) = \ln(\beta) + \ln(x)$
- Intercept term [$\ln(\beta)$] expected to be log of proportionality factor
  - Alternatively, the antilog of intercept is the proportionality factor
- Slope [coefficient of $\ln(x)$] expected to be one

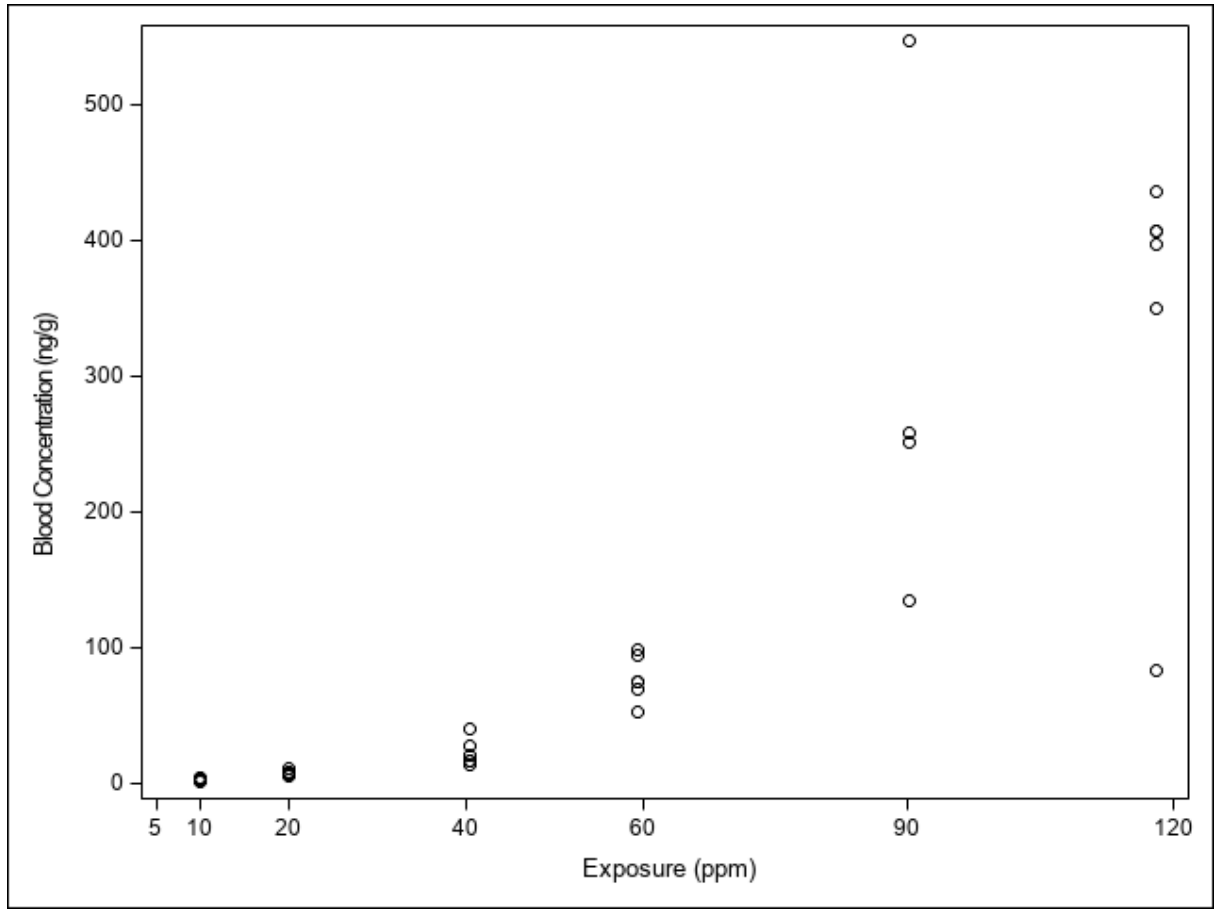
Proportional Relationship – Original Scale

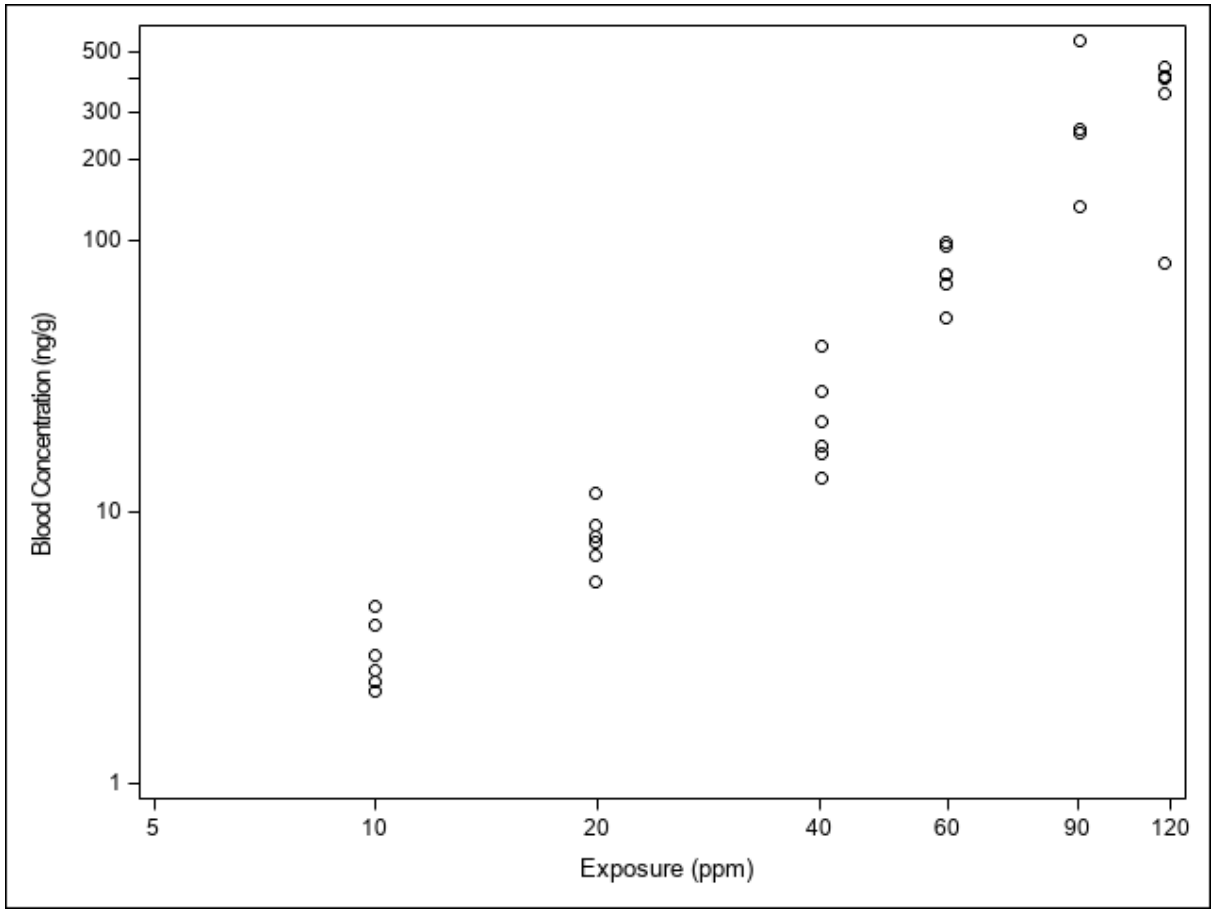


Proportional Relationship – Log-Log Scale

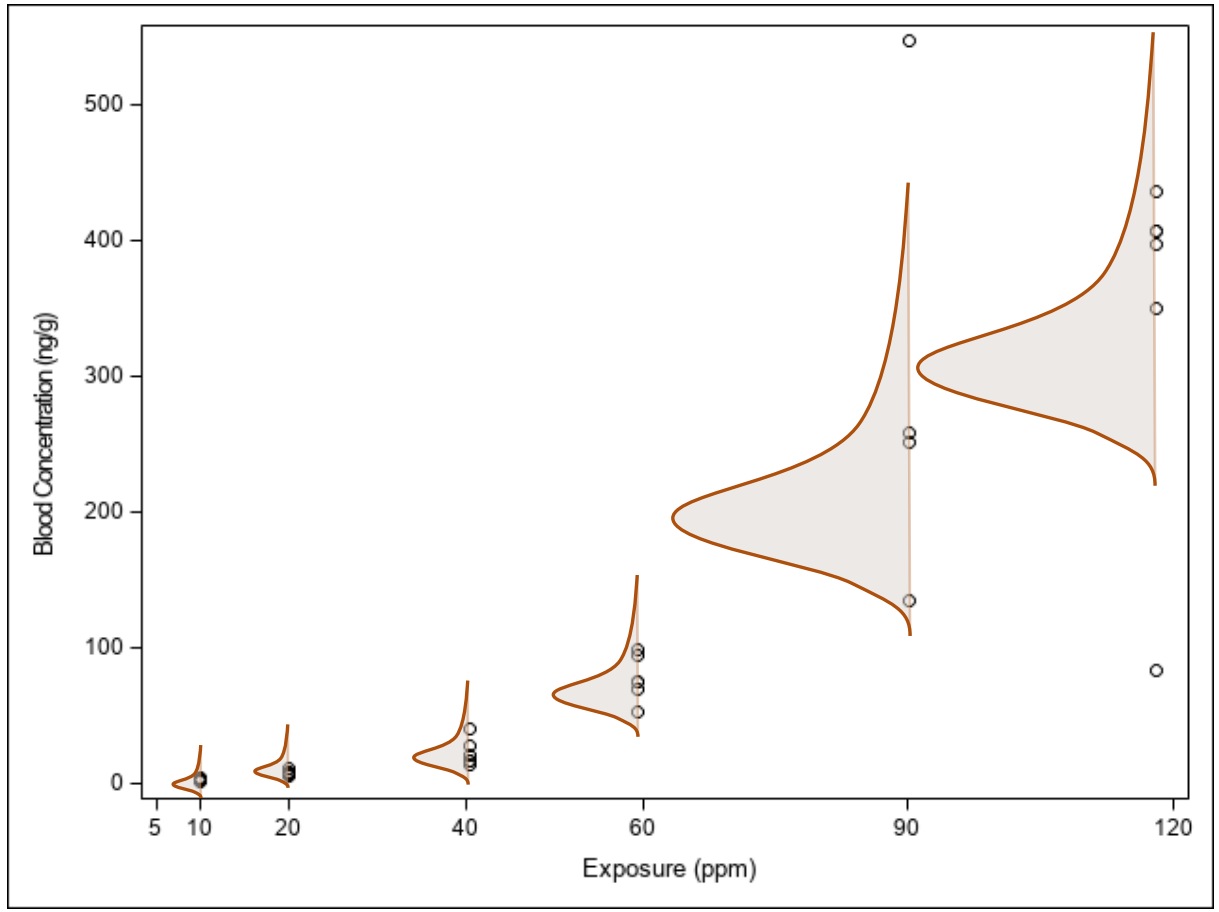


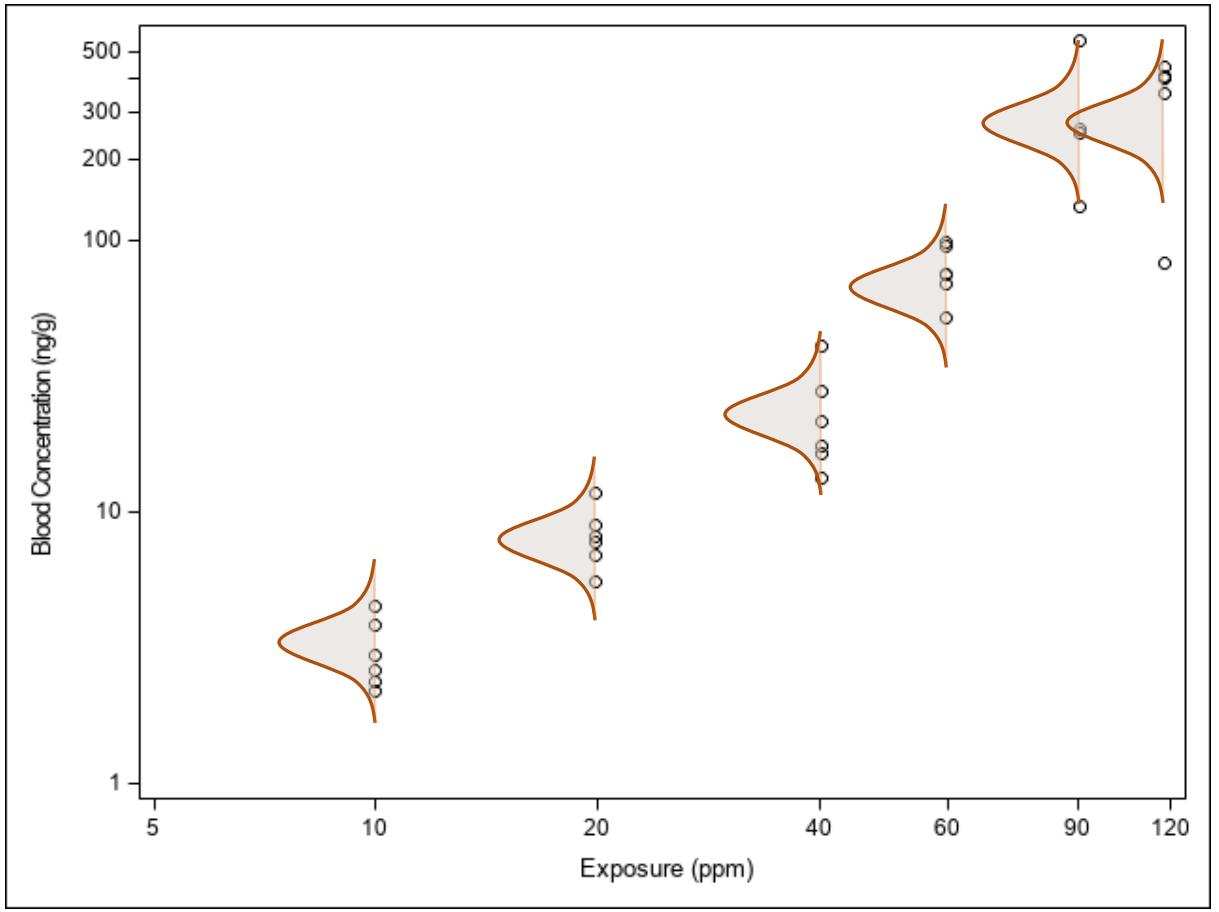
Actual Concentration in Blood vs. Exposure Dose



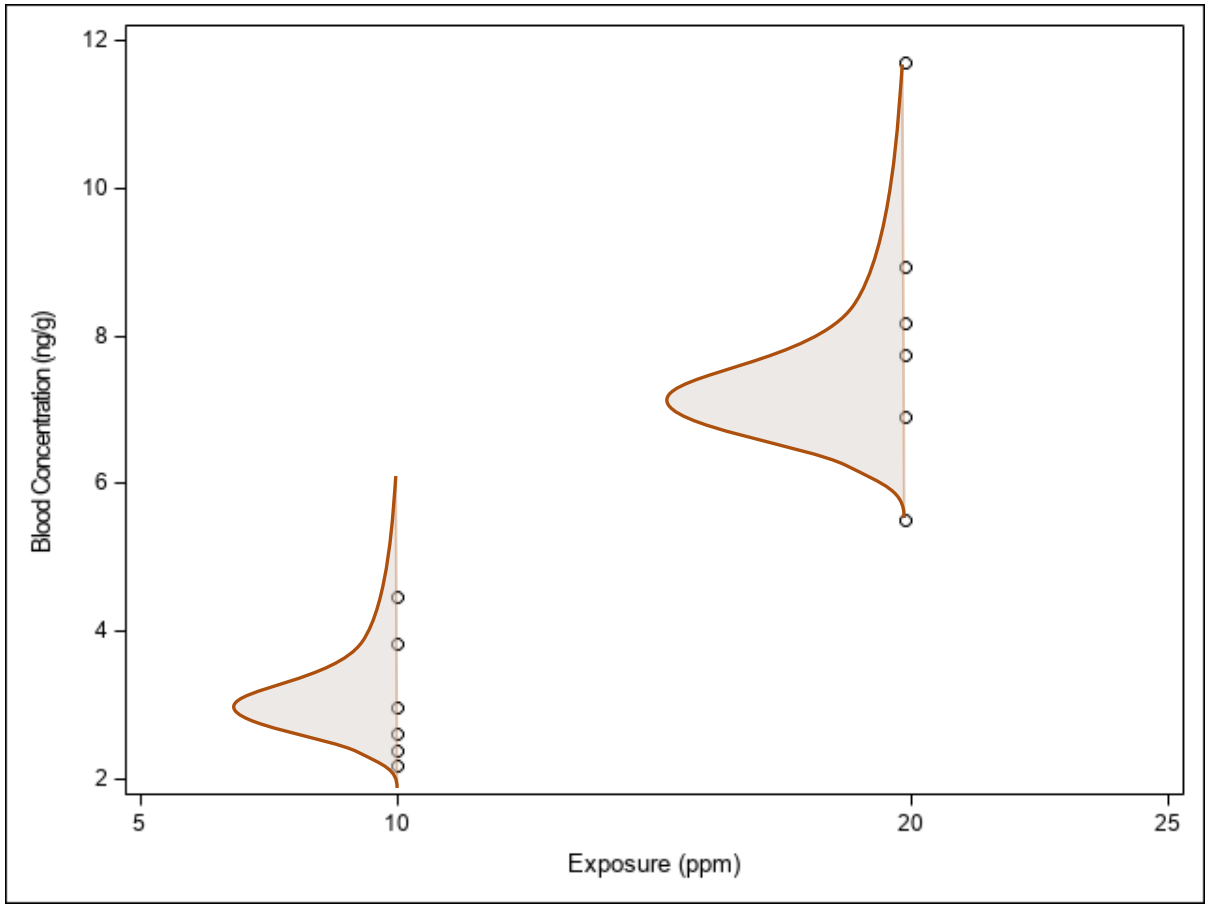


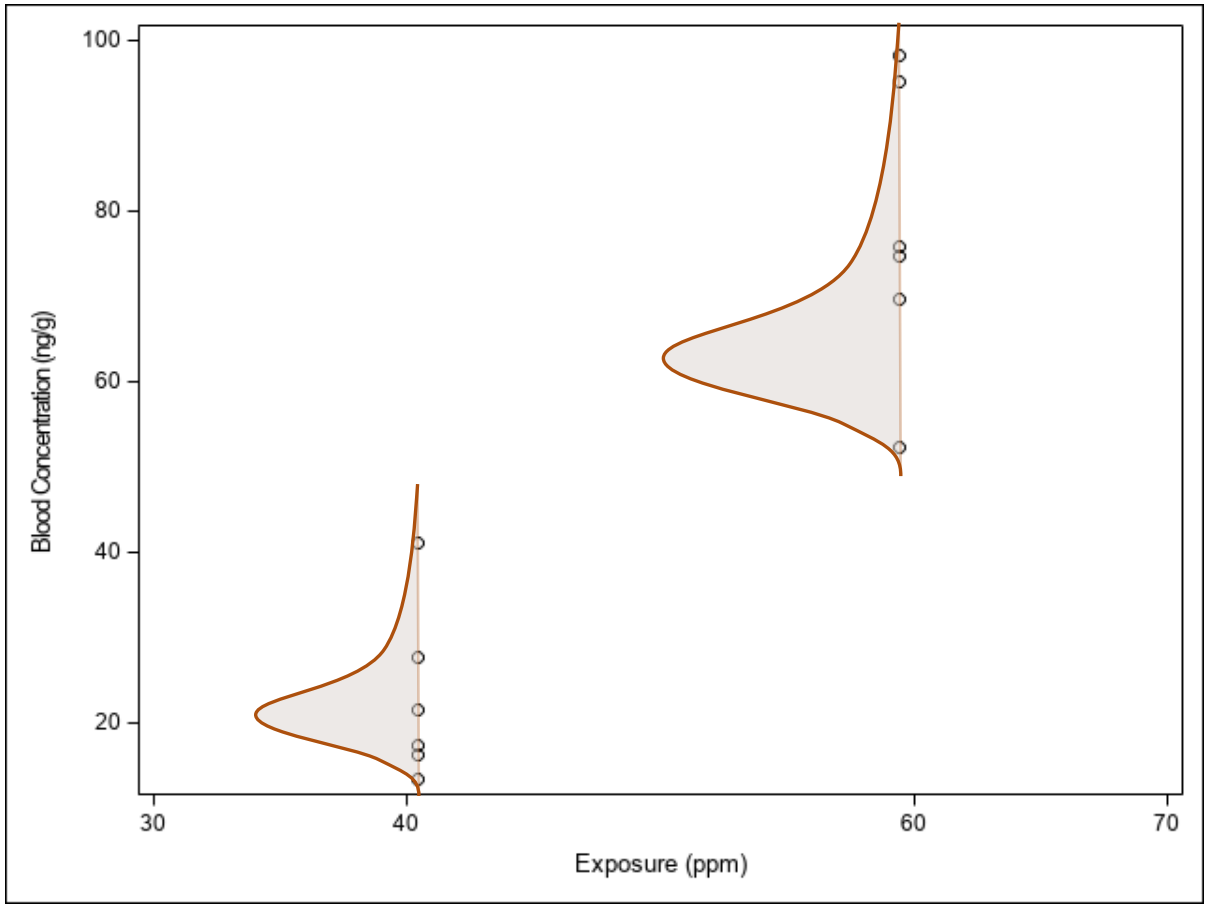
Actual Concentration in Blood vs. Exposure Dose





Actual Concentration in Blood vs. Exposure Dose - Low & Mid Doses





Comments Case-Study Data

Distributed Lognormally

- Normal distribution appears not to be good fit for un-transformed data
- Concentration data typically follow lognormal distributions
- Concentration in blood appear to be lognormally distributed

Exhibits heterogeneity of variance (unequal variance)

- Variance of blood concentration is greater as measurements increase

Limited options of statistical models to analyze un-transformed data

- Due to severe violations of standard assumptions (heterogeneity of variance assumption and or normality)

Log-Transformation Needed

- Meets assumptions of normality and homogeneity of variance for regression analysis
- Results in asymmetrical confidence intervals more appropriate to data

Important Considerations for Regression

Toxicologist input is needed to determine if conceptual model is appropriate

- Biological relevance and interpretation of parameter estimates are critical

Model should be fit to individual observations, not just group means

- Parameter estimates and confidence intervals will account for observed variability

Complexity of model is limited by number of doses groups

- More complex models generally have more parameters
- There should not be more parameters than dose groups
- Minimum number of dose groups needed to splice 2 linear models

Log-Log Regression

Using Gough 1995 model, we can express the blood concentration vs. exposure dose of subject an i^{th} as:

$$y_i = a \times exposure^b \times e^{\varepsilon i} \qquad \text{(eq. 1)}$$

where a and b are constants and can be estimated from the data, and ε_i is i.i.d. and $\varepsilon_i \sim N(0, \sigma^2)$

Taking log of both sides of the equation above, we have

Model 1: $\log(y_i) = \log(a) + b \times \log(exposure) + \varepsilon_i$

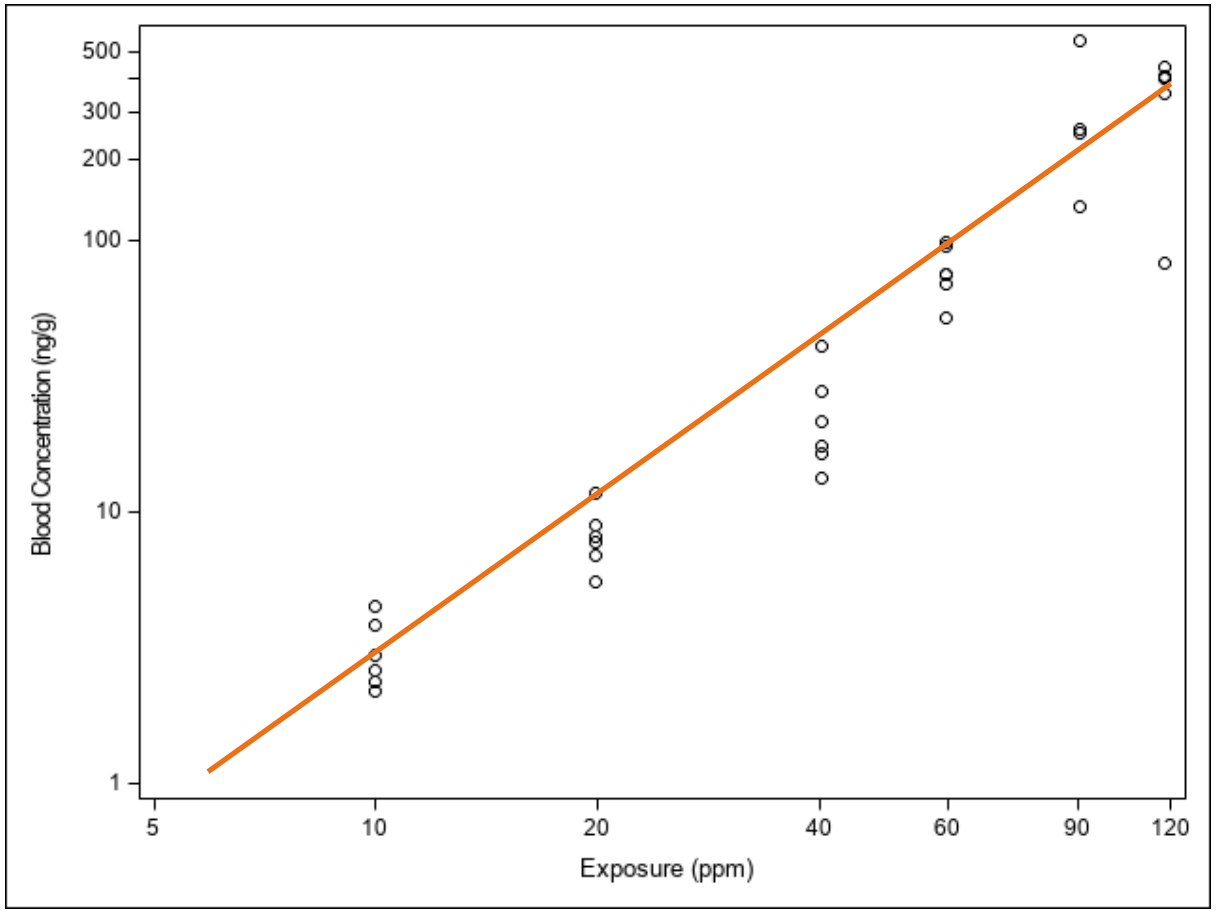
If slope $b = 1$ (or not statistically different from 1), blood concentration is proportional (or reasonably assumed proportional) to exposure for whole range of exposure

If the slope b significantly $\neq 1$, there is evidence blood concentration is not proportional to exposure for entire range of exposure

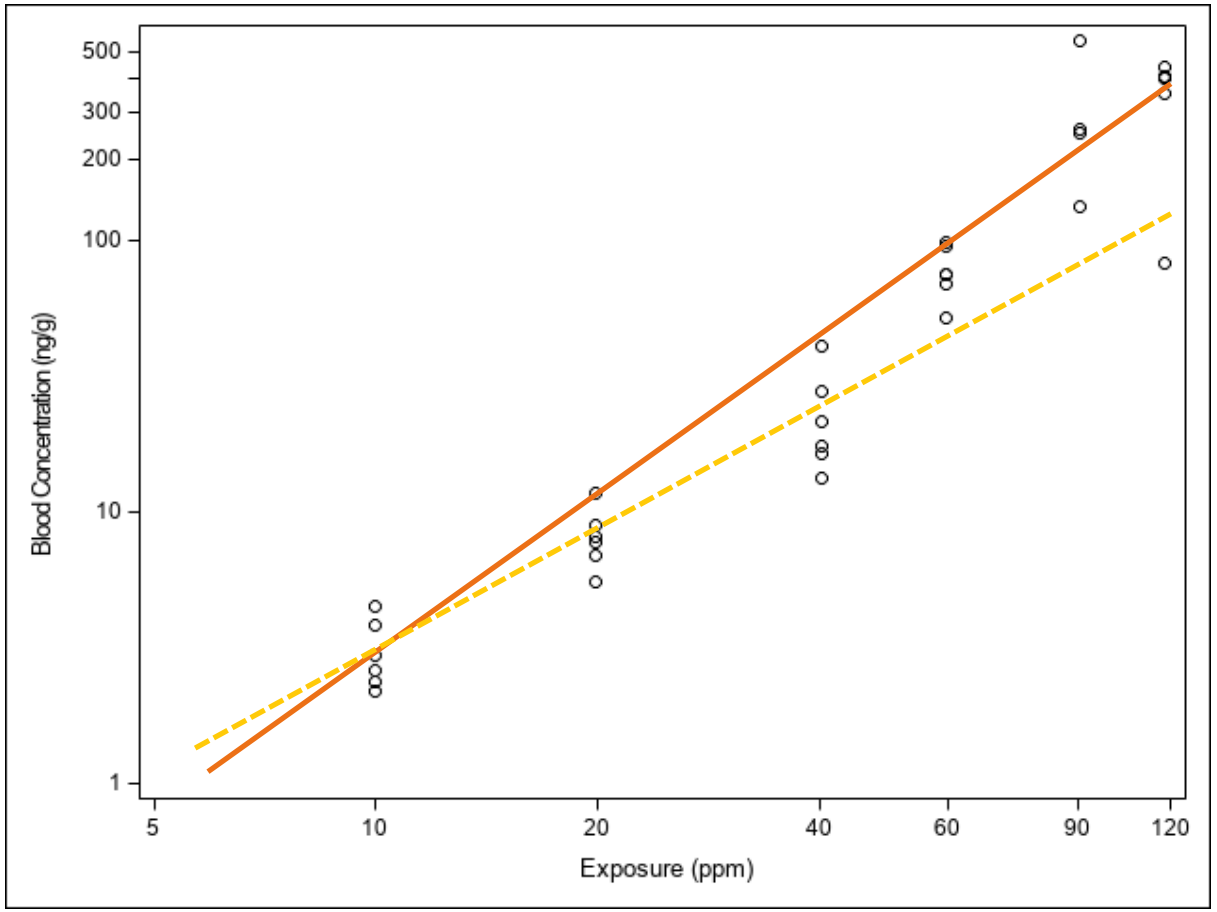
Regardless whether $b = 1$ or not, a Lack-of-Fit F-test will be performed to determine whether Model 1 adequately fits data

- If there is no evidence Model 1 does not adequately fit data (p-value > 0.05), then accept results of Model 1
- If there is evidence Model 1 does not adequately fit data (p-value $\leq$ 0.05), then a single slope b in Model 1 is probably not adequate to characterize relationship between blood concentration vs. exposure for entire range of exposure

Testing Slope & Fit of Log-Log Regression

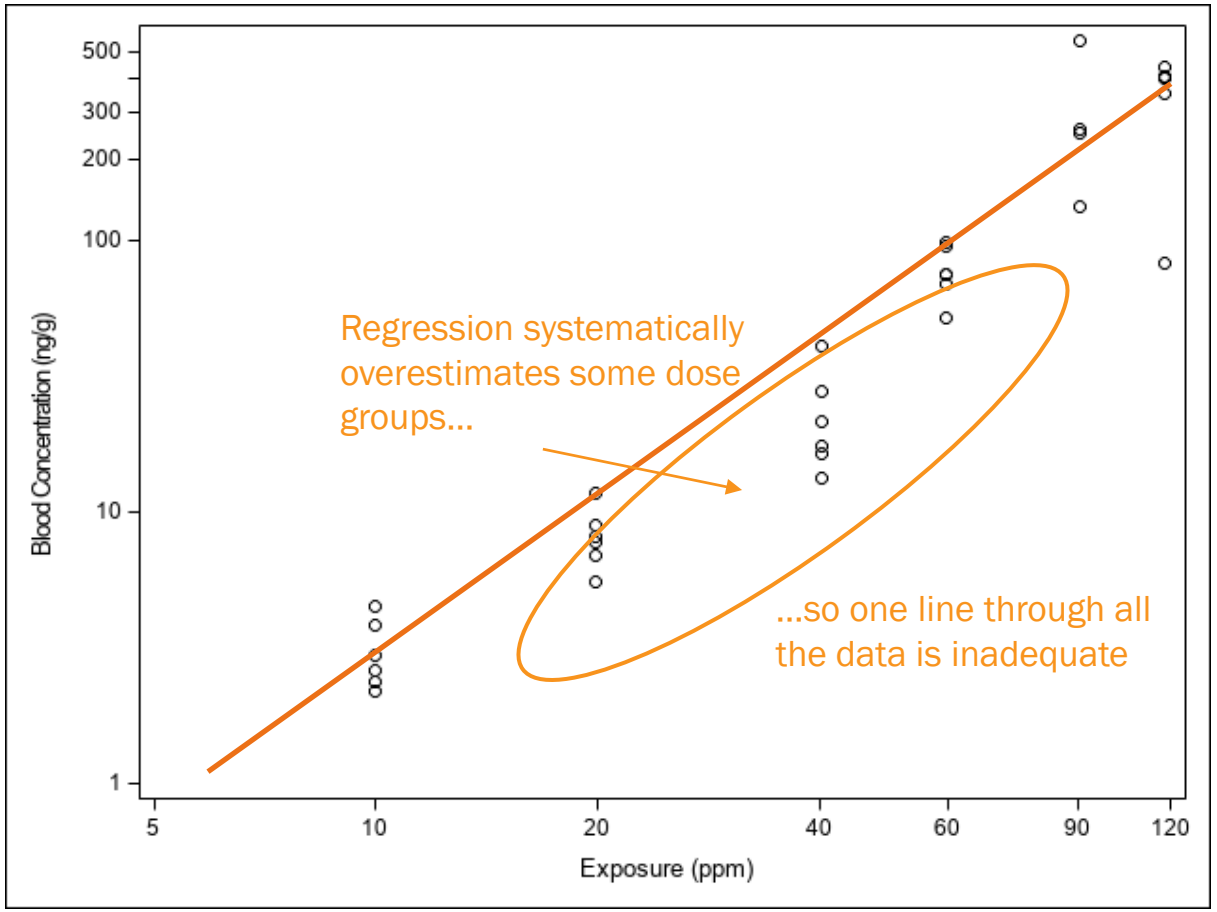


Testing Slope & Fit of Log-Log Regression



Does the slope from the simple log-log regression differ significantly from 1 (yellow dashed line)?

Testing Slope & Fit of Log-Log Regression



Does the slope from the simple log-log regression differ significantly from 1 (yellow dashed line)?

Does the lack of fit test indicate the model does not adequately fit the data?

Piecewise Regression

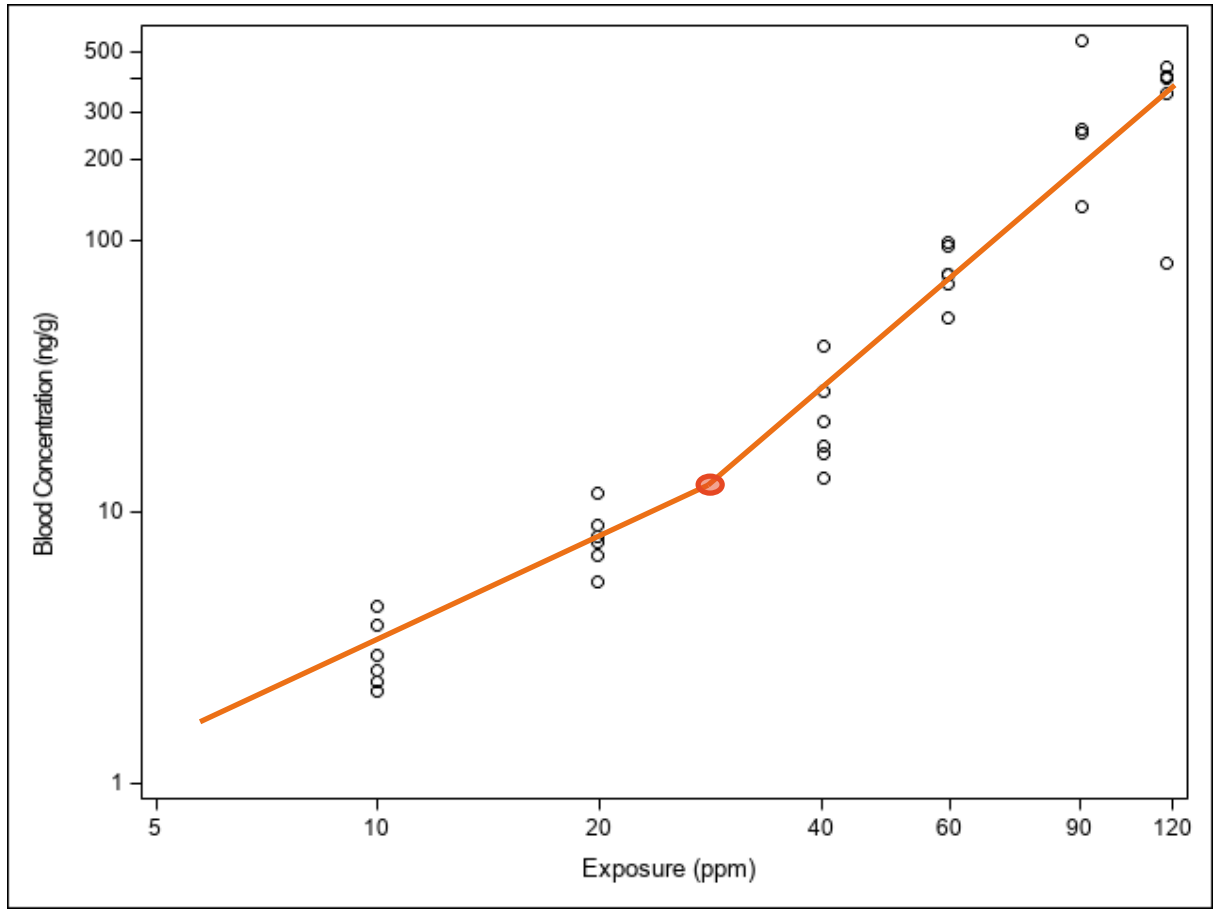
When Lack-of-Fit F-test indicates Model 1 doesn't adequately characterize relationship between blood concentration and exposure for entire range of exposure then:

- Assume relationship between blood concentration and exposure changes at X_0
- The relationship between blood concentration vs. exposure is characterized by Model 2 below

Model 2

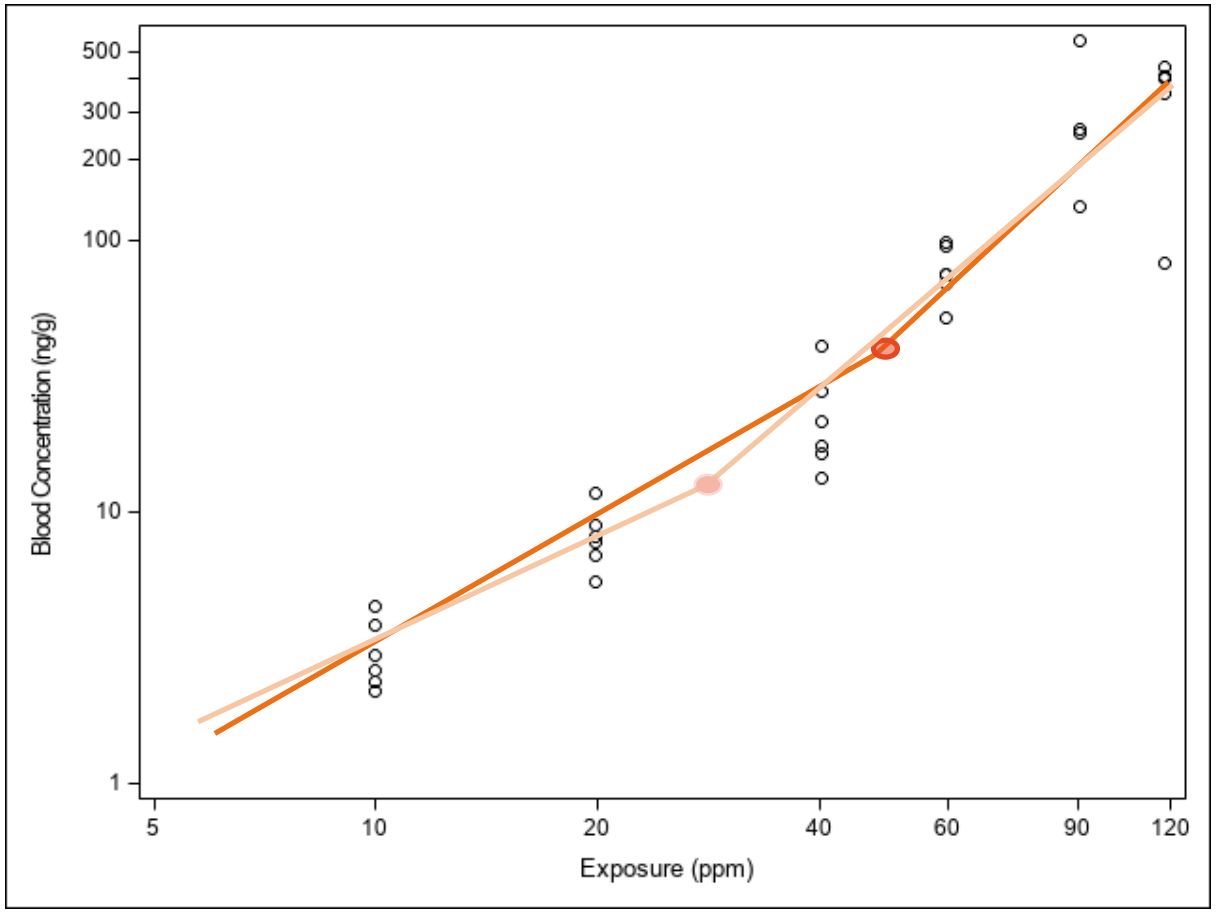
$$\log(y_i) = \begin{cases} \log(a) + b \times \log(exposure) + \varepsilon_i & \text{if } exposure \leq X_0 \\ \log(a) - \Delta b \times \log(X_0) + (b + \Delta b) \times \log(exposure) + \varepsilon_i & \text{if } exposure > X_0 \end{cases}$$

Determining Significant Departure from Proportionality



Model fits two regression lines "spliced" at a knot, the estimated KMD

Determining Significant Departure from Proportionality



Model fits two regression lines "spliced" at a knot, the estimated KMD

Where should we place the knot?

Case Study Example: Model 1

Regression Model 1: $\log(y_i) = \log(a) + b \times \log(exposure) + \varepsilon_i$

or equivalently expressed in eq. 1: $y_i = a \times exposure^b \times e^{\varepsilon i}$

Parameter	Parameter Estimate	Approx Standard Error	Approximate 95% Confidence Limits	
log(*a*)	-3.691	0.368	-4.441	-2.941
b	1.962	0.097	1.764	2.161

Estimated slope *b* = 1.962 (95% CI = 1.764 – 2.161) is significantly different from 1 → **blood concentration is not proportional to exposure over entire range of exposure**

However, Lack-of-Fit F-test indicates Model 1 **inadequately characterizes relationship** between blood concentration and exposure (p-value = 0.0112)

- The single straight line should not be used to fit data

Case Study Example: Model 2

Regression Model 2:

$$\log(y_i) = \begin{cases} \log(a) + b \times \log(exposure) + \varepsilon_i & \text{if } exposure \leq X_o \\ \log(a) - \Delta b \times \log(X_0) + (b + \Delta b) \times \log(exposure) + \varepsilon_i & \text{if } exposure > X_o \end{cases}$$

or equivalent expressed:

$$y_i = \begin{cases} a \times exposure^b \times e^{\varepsilon_i} & if\ exposure \leq X_0 \\ a \times \frac{1}{X_0^{\Delta b}} \times exposure^{b+\Delta b} \times e^{\varepsilon_i} & if\ exposure > X_0 \end{cases}$$

Case Study Example: Model 2

Parameter	Parameter Estimate	Approx Standard Error	Approximate 95% Confidence Limits	
$\log X_0$	3.5886	0.3651	2.8429	4.3344
$\log(a)$	-2.2004	0.9829	-4.2077	-0.1930
b	1.4286	0.3683	0.6765	2.1807
Δb	1.1173	0.4306	0.2378	1.9968

Slope *b* changes at approximately exposure = $e^{3.5889}$ = 36.18 ppm (95% CI = 17.17 - 76.28)

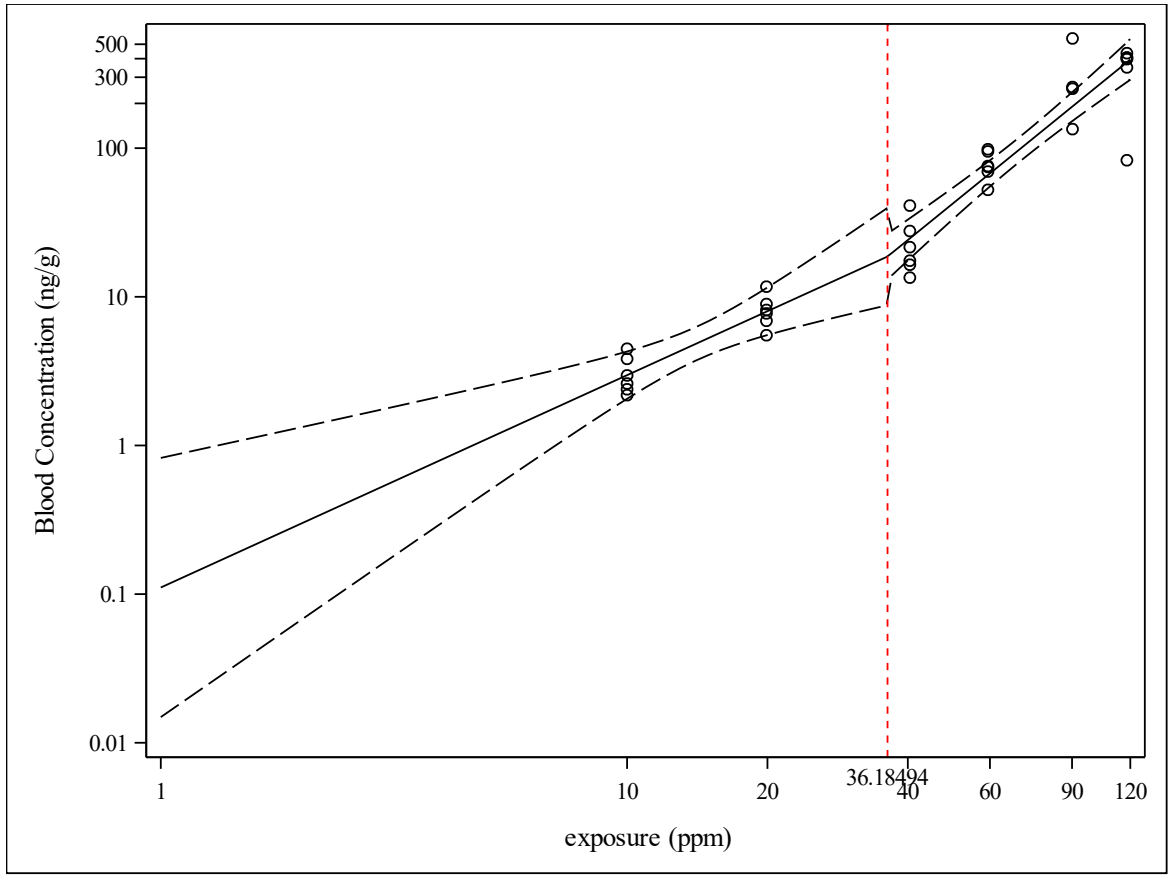
When exposure ≤ 36.18 ppm, estimated slope *b* = 1.4286 (95% CI = 0.6765 – 2.1807) is not significantly different from 1 → **relationship is not significantly different from proportionality**

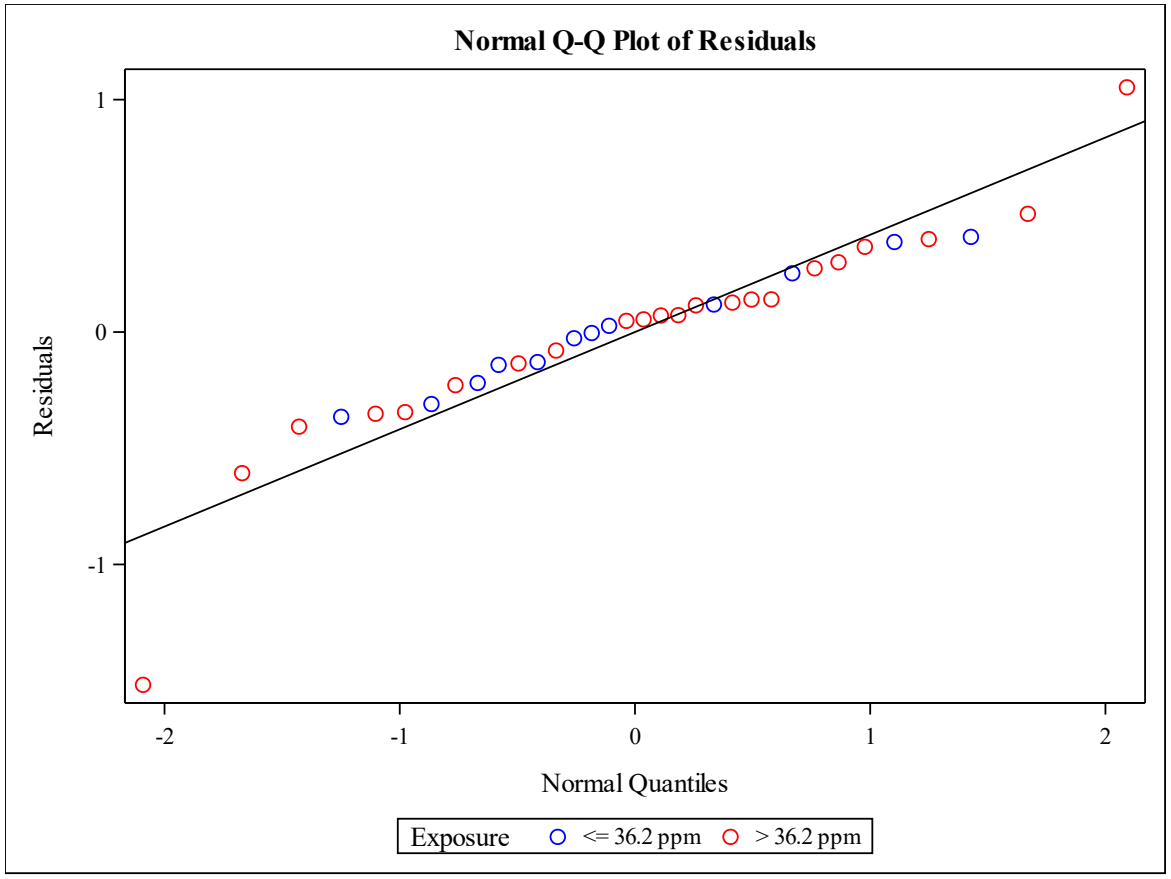
Change in slope at exposure = 36.18 ppm is significant Δb = 1.1173 (95% CI = 0.2378 – 1.9968) → **Dose at which slope significantly departs from approximately proportional**

When exposure > 36.18 ppm, estimated slope $b + \Delta b$ = 2.5459 (95% CI = 2.0900 – 3.0017) is significantly different from 1 → **relationship is significantly more than proportional**

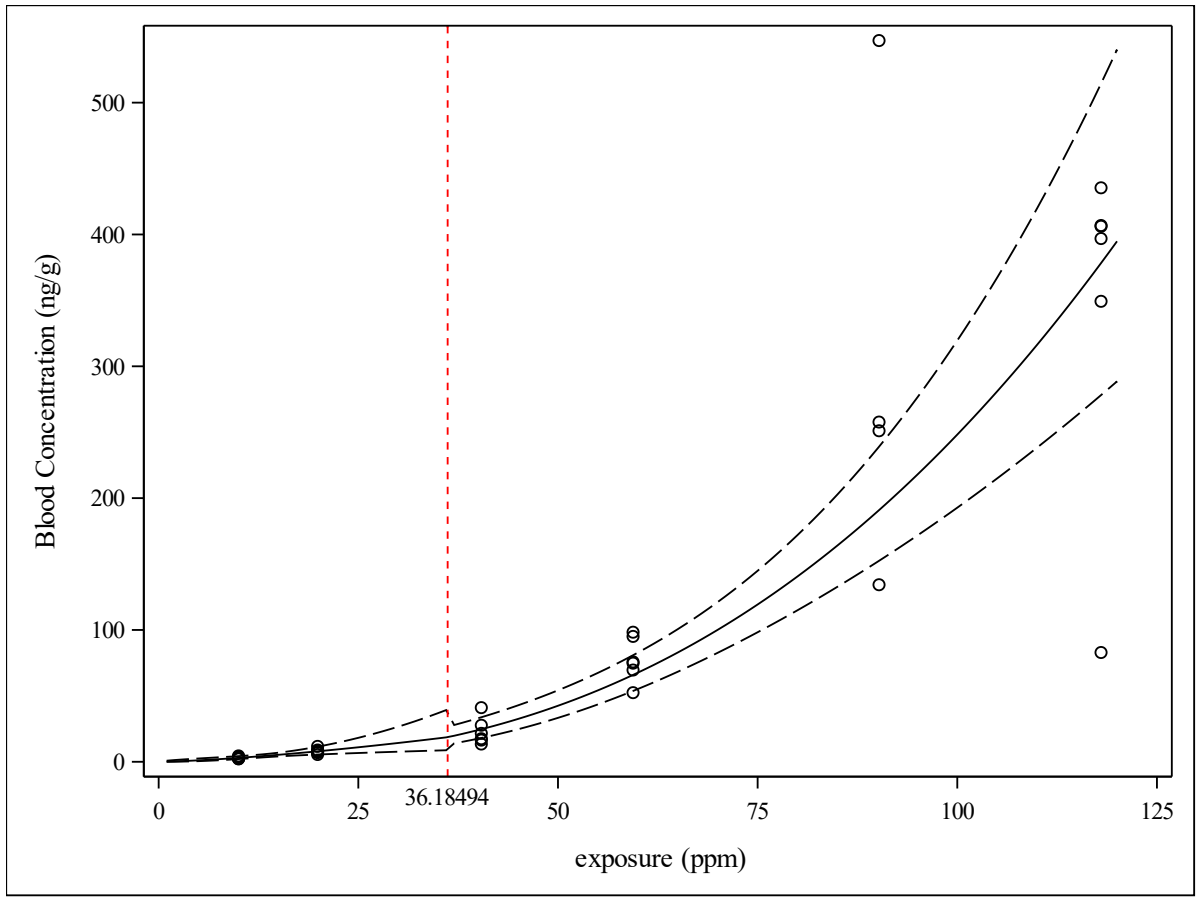
Lack-of-Fit F-test results in p-value = 0.0876, indicates Model 2 **adequately characterizes data**

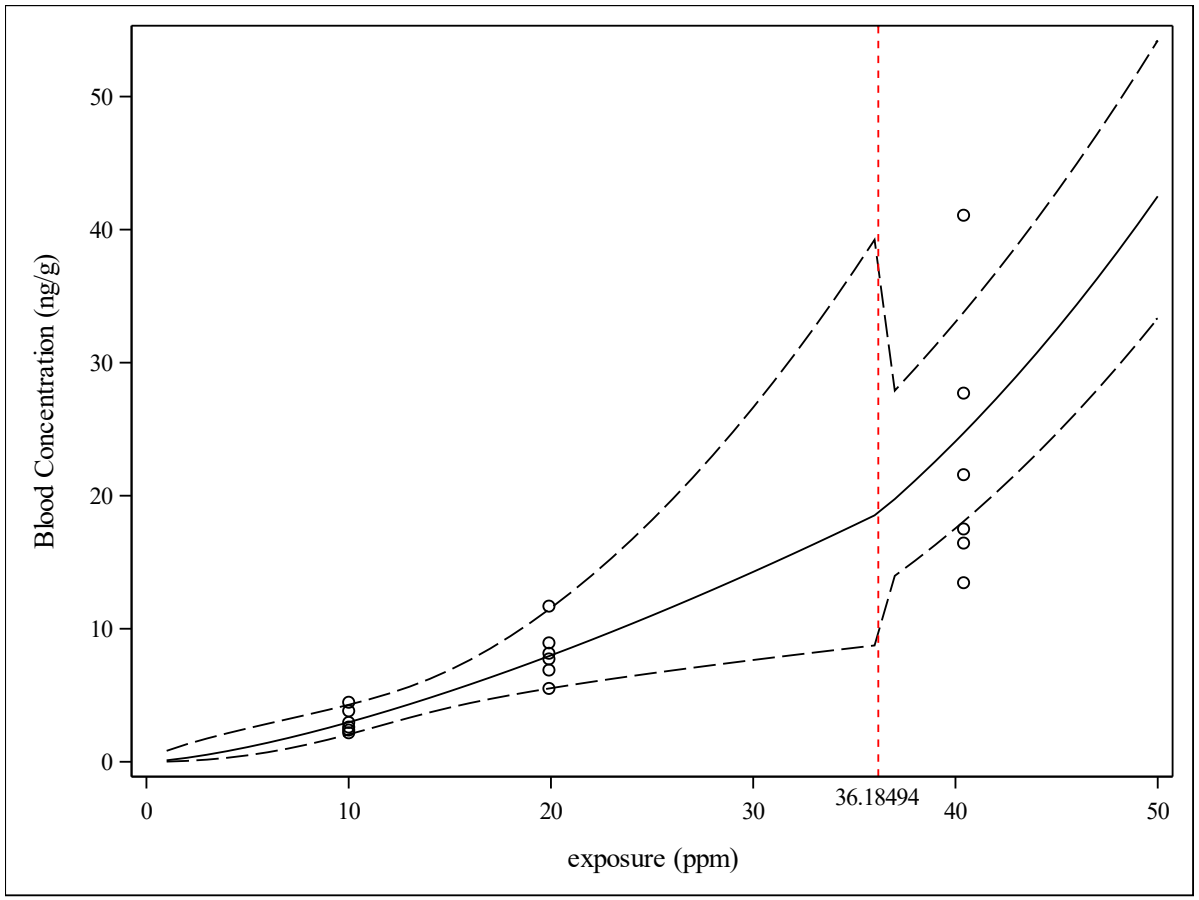
Model 2 – Log-Log Scale





Model 2 – Original Scale





Comparing Models

Model 2 vs. Model 1

- F-test was used to compare models
  - A rigorous, mathematical and objective way to select an appropriate model
- Result indicates Model 2 was significantly better than Model 1 to characterize relationship between blood concentration vs. exposure (p-value = 0.016)
  - The piecewise regression model with a knot is a better fit than a single straight line

Model 1			Model 2			F-test			
SSE	p	MSE	SSE	p	MSE	F-value	DF1	DF2	p-value
7.616	2	0.238	5.780	4	0.193	4.764	2	30	0.016

Comparison to Other Approaches

Piecewise Regression

- Incorporates individual observations & observed variability
- Fits relationship using all dose groups
- Provides statistical tests to determine a significant departure from proportionality

Comparing fold differences

- Uses only group means and does not account for variability within dose groups

Estimating linear relationship between (0,0) and first data point

- Uses only group means and does not account for variability within dose groups
- Fits relationship based on one dose group

Conclusions

Toxicologists provide critical insight into biological relevance and plausibility of any models being fit

Statisticians can translate questions about data characteristics into mathematical and testable statements

- e.g., where does the relationship between internal and external dose significantly depart from proportionality?

Data should be appropriately transformed to meet any underlying assumptions of statistical analysis

- Relationship is approximately linear, variance are normal and heterogeneous, etc.

Any statistical analysis should attempt to incorporate all dose groups and individual observations to appropriately characterize the variability and modeled relationship

Statistic methods can

- Quantitatively address uncertainty in KMD estimates using confidence bounds
- Estimate KMDs that may exist *between* dose groups, rather than being limited to selecting a tested dose group

Acknowledgements

Thank you to **David J. Miller** & **Cecilia Tan** for thoughtful review & suggestions

Special appreciation to **James Nguyen** for his statistical expertise and programming skill!