Comparing methods for averaging IC50s: can we forego complexity for a simpler approach?

Kristen Kohler,[1] Traymon Beavers,[1]
Paul Dudas,[2] John DeLong,[2]
Clara Moon,[2] Kacey Sachen,[2]
Indra Sarabia,[2] Jocelyn Sendecki [1]

NCS Conference 2024
September 25
Wiesbaden, Germany

Johnson&Johnson
Innovative Medicine

[1] Johnson & Johnson US, Discovery and Nonclinical Safety Statistics; [2] Johnson & Johnson US, Discovery Immunology

Agenda

1. Background & Motivation
2. Methods for Aggregating IC50s
3. Comparing Methods via Simulation
4. An Automated Solution
5. Conclusions & Next Steps

Background & Motivation

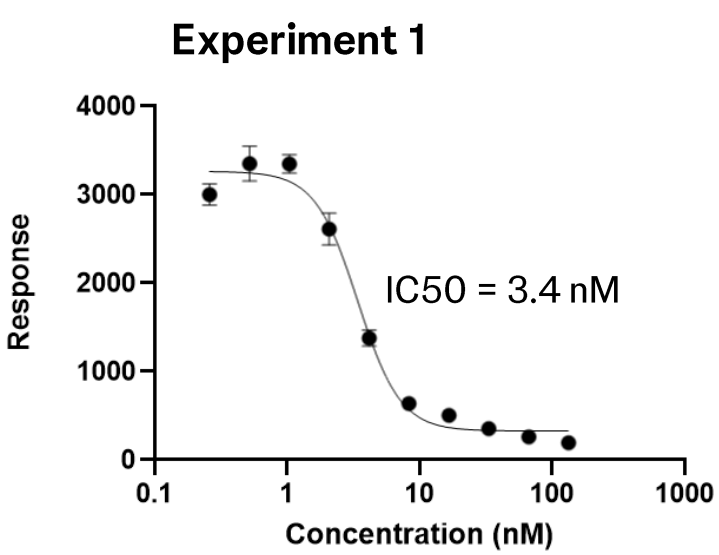
Motivation

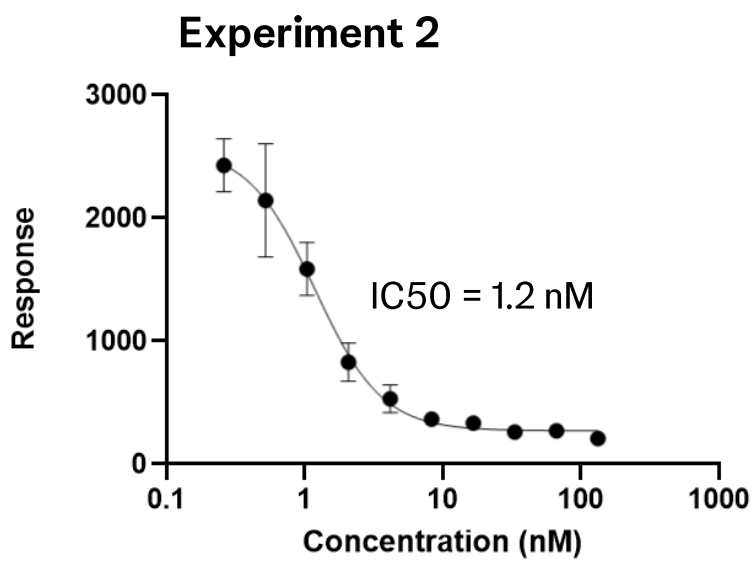
- First in-human dose of a treatment is established in part by performing repeated concentration-response experiments and determining IC50 (concentration corresponding with half maximal response)
- Many scientists have contacted us regarding averaging IC50s for various discovery programs
- Need a sound way to aggregate IC50s across experiments or donors
  - Statistically accurate
  - Easy to understand

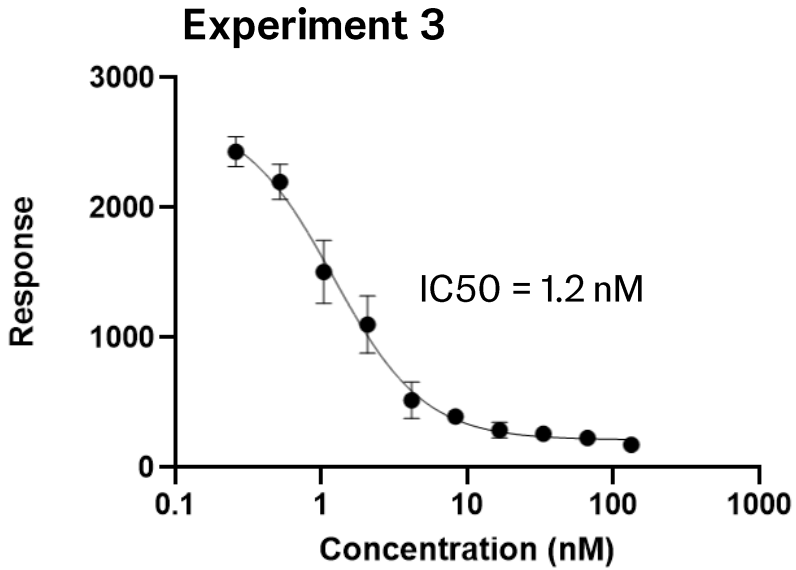
Motivation – A Practical Example

Example 1: Inhibition Data (Species 1) for Compound A

- Asked to average IC50s for small number of *in vitro* experiments for inhibitor of pro-inflammatory cytokine







Motivation – A Practical Example

Example 1: Inhibition Data (Species 1) for Compound A

- Original Method: Calculate arithmetic average in excel
  - Over-predicts average since IC50 is log-normally distributed

Experiment	IC50 (nM)
Experiment 1	3.4
Experiment 2	1.2
Experiment 3	1.2

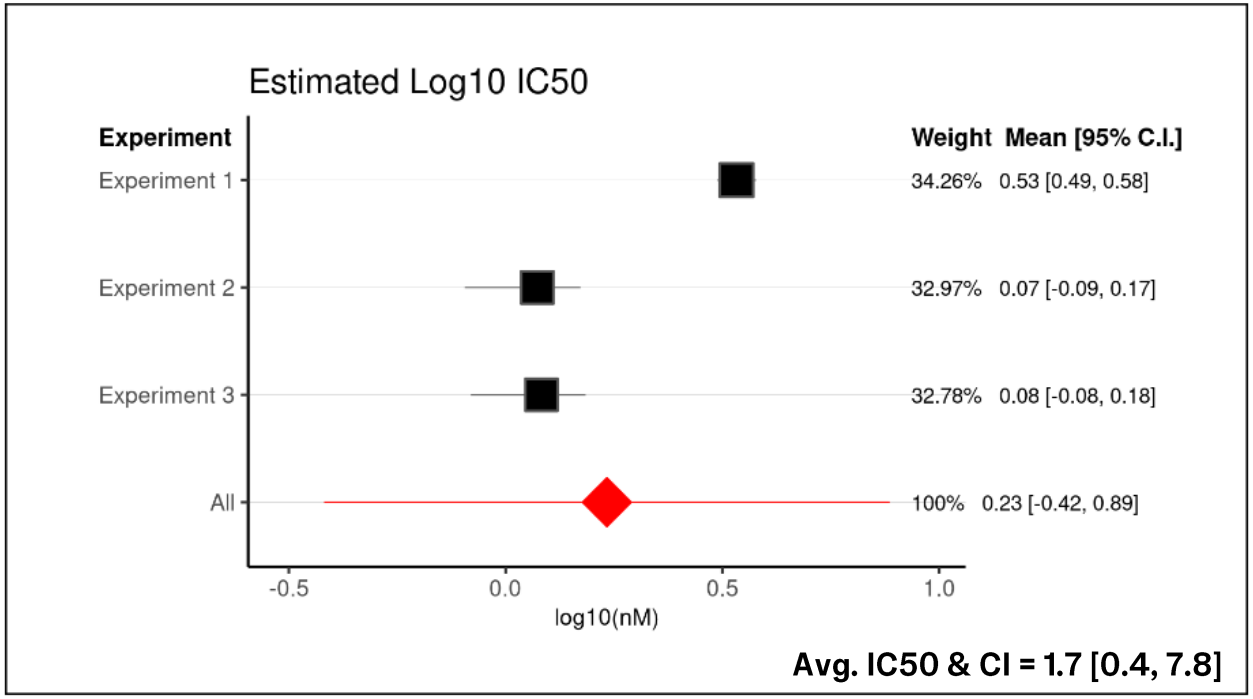
Measure	Estimate (nM)
Arithmetic Mean	1.9
Margin of Error	1.5

Average IC50 & MoE
1.9 nM +/- 1.5 nM

Motivation – A Practical Example

Example 1: Inhibition Data (Species 1) for Compound A

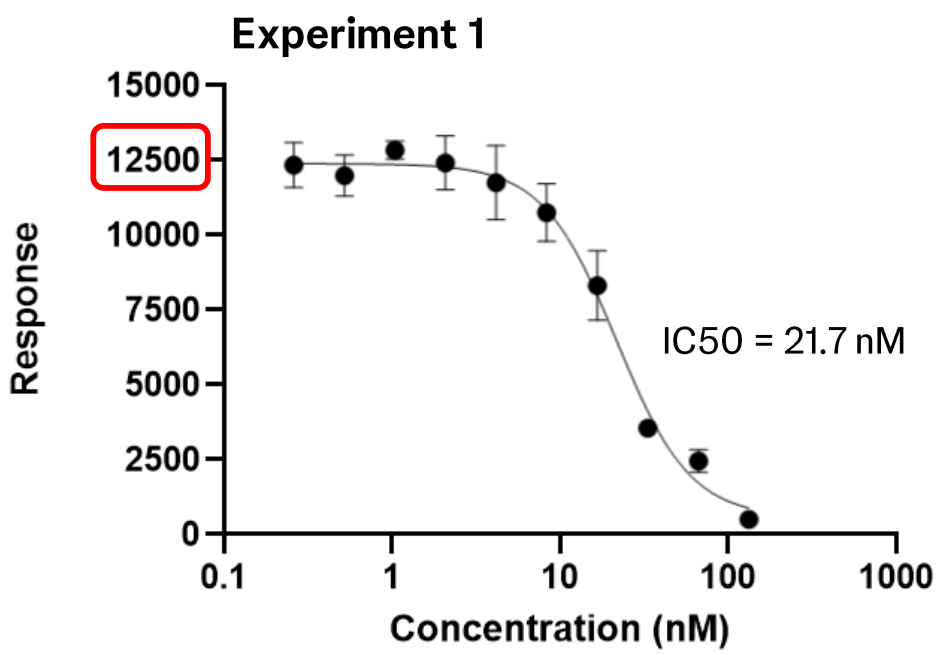
- Perform fixed-effect meta-analysis with Hedges' estimator and t-test confidence intervals by recommendation of Jiang et al. *
- Confidence intervals 'inflated' with t-test for small number of experiments due to low degrees of freedom

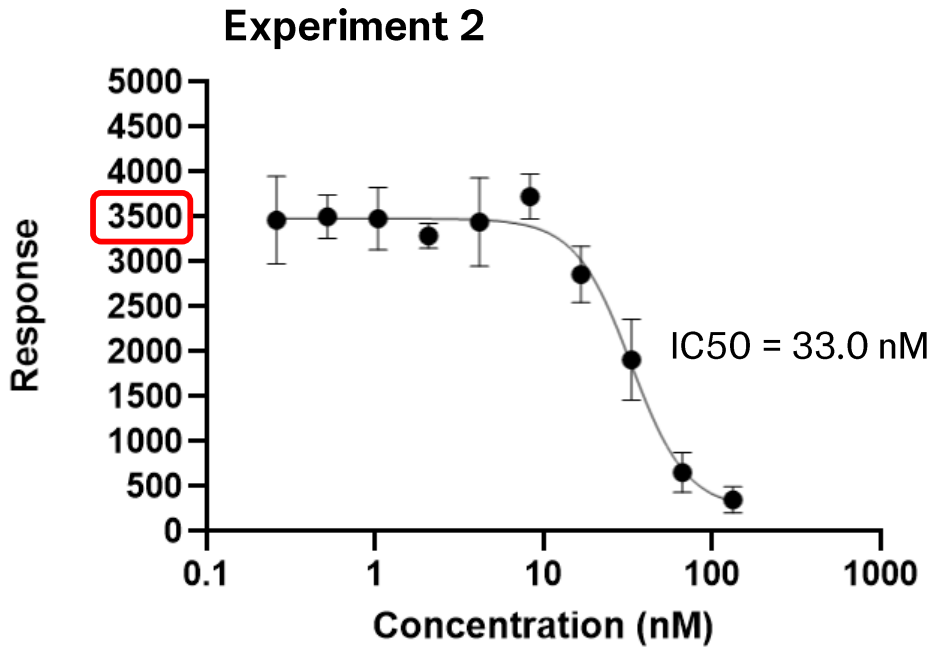


* '*Summarizing EC50 Estimates from multiple dose-response experiments: A comparison of a meta-analysis strategy to a mixed-effects model approach*', Jiang X. , Kopp-Schneider A.

Motivation – A Practical Example

Example 2: Inhibition Data (Species 2) for Compound A – Different Scales

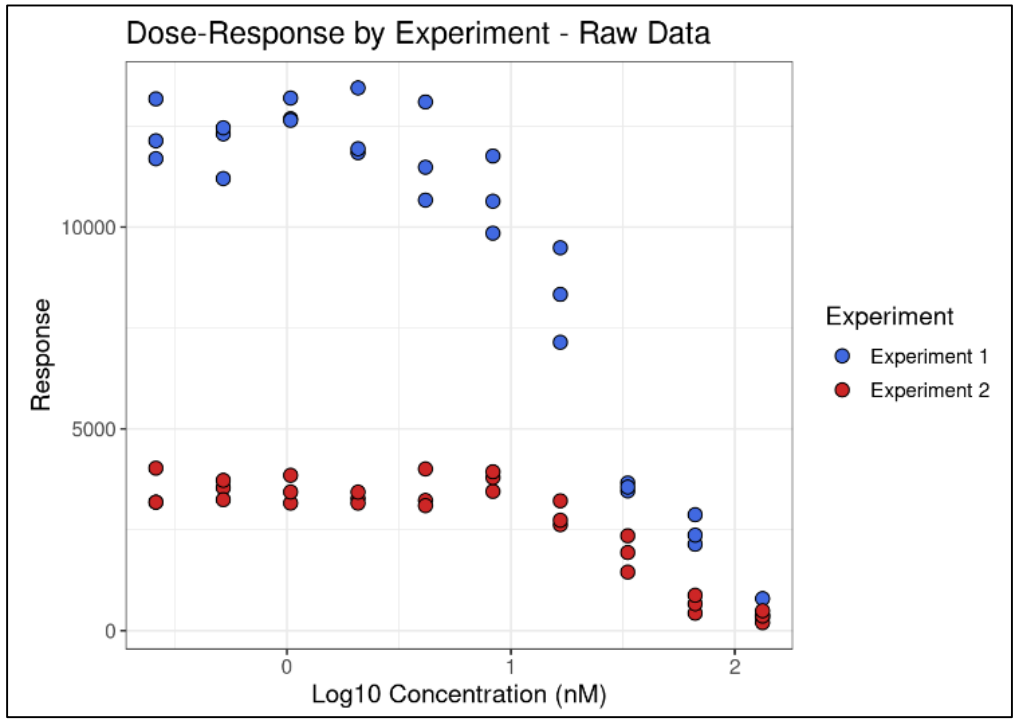


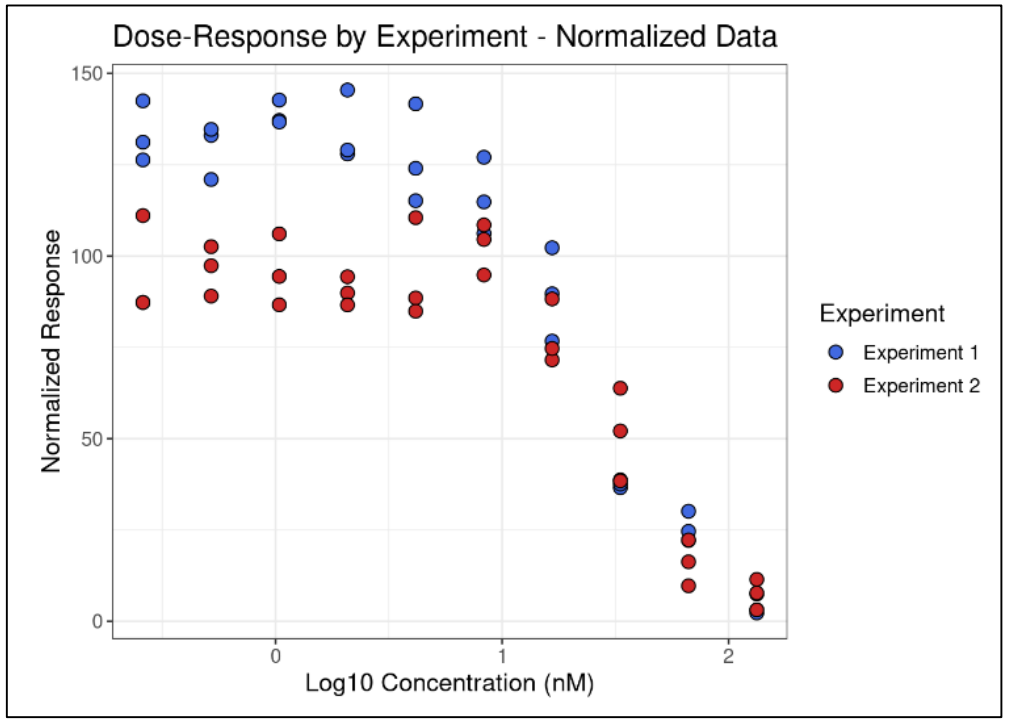


Motivation – A Practical Example

Example 2: Inhibition Data (Species 2) for Compound A – Different Scales

- Non-linear mixed effect (NLME) model recommended for N = 2*
- When data is on very different scales, transformation may be required





* *'Summarizing EC50 Estimates from multiple dose-response experiments: A comparison of a meta-analysis strategy to a mixed-effects model approach'*, Jiang X. , Kopp-Schneider A.

Goals

- Find a method that is statistically accurate and simple
- Determine whether complex approaches (NLME) yield results similar enough to simple approaches (meta-analysis) to use a simplified approach moving forward
- Create an automated solution for scientists to calculate average IC50s

Methods for Aggregating IC50s

Geometric Mean

Report geometric mean across experiments as average IC50

Advantages

- Easy to understand
- More accurately predicts average IC50 than an arithmetic mean
- Can be calculated without help from statistician

Limitations

- Does not account for the variance associated with the estimated IC50s

Meta-analysis

Average IC50 calculated by weighting individual Log IC50 estimates by standard error

Advantages

- Easy to understand
- Fixed or random effects
- Can be implemented via Excel or Shiny Application
- Can use estimates from GraphPad Prism as input

Limitations

- Not recommended for less than three experiments*
  - Common in Discovery due to cost
- Confidence intervals could be 'inflated' with small number of experiments
- Could lose information by weighting curves with high standard error to zero

* *'Summarizing EC50 Estimates from multiple dose-response experiments: A comparison of a meta-analysis strategy to a mixed-effects model approach'*, Jiang X. , Kopp-Schneider A.

Non-Linear Mixed Effect (NLME) Model

Model data for all experiments/donors together to calculate an aggregate IC50

Advantages

- Recommended method for calculating average IC50
- Works well when averaging small number of experiments

Limitations

- Convergence issues
- Results are misleading if curves do not have similar min and max parameters
  - May require transformation of data
- More complex – may require assistance of statistician

Examples Through Simulation - Methods

Methods for Comparison

1. NLME Models
   - Performed on raw data and normalized data
   - Explored with random effects on all parameters and random effect only on IC50

2. Meta-analysis
   - Performed on individual experiment Log IC50 estimates, normalization not needed
   - Explore random effects and fixed effects
     - DL Estimator (random effects)
     - Hedges' Estimator (fixed effects)
   - Both methods explored with z-test and t-test confidence intervals

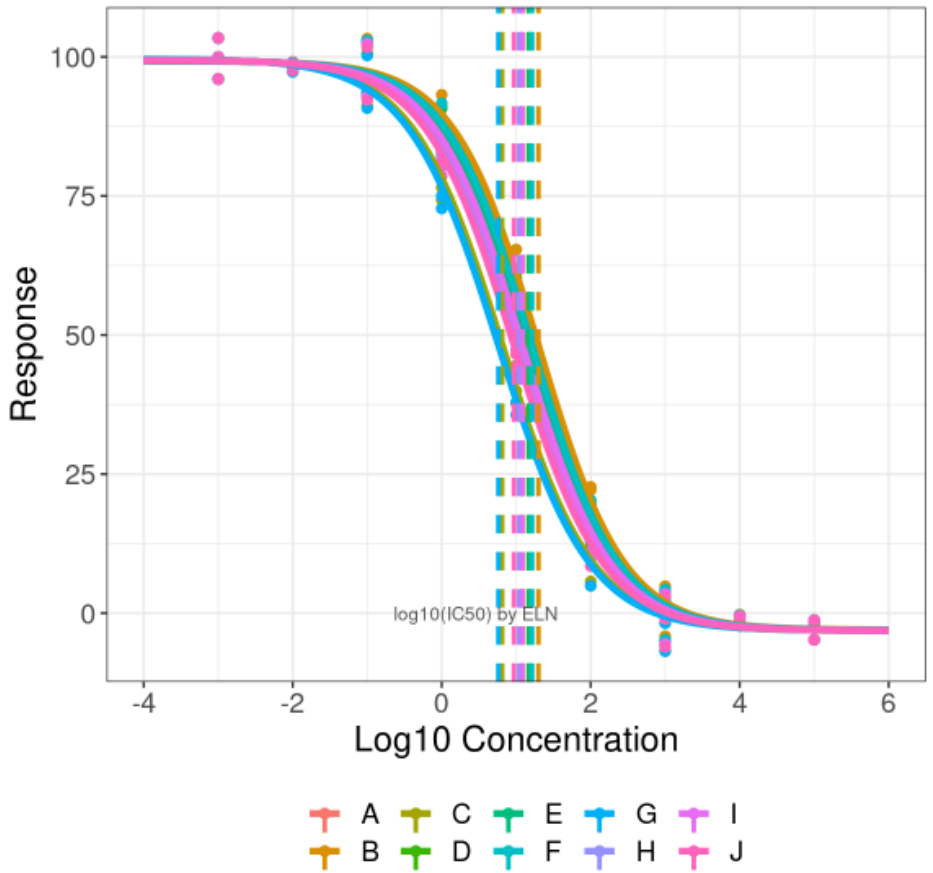
Simulation Process

- Simulate 10 dose-response curves to replicate experiments
- Compare average IC50s & confidence intervals across methods when averaging 2 experiments and 10 experiments under two scenarios:
  - Stable case (e.g., cell lines)
    - Very similar dose-response curves between experiments
  - Curves with more variation and on different scales (e.g., animal data)
    - Different top parameters

Examples Through Simulation – Low Variation

Simulation Setup – Low Variation Between Curves

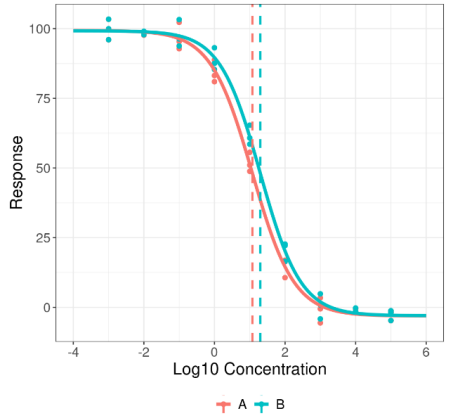
- Simulated dose-response curves for 10 experiments
- Low within-experiment variation for all experiments
- Little variation between IC50 estimates
- Nearly identical top and bottom parameters

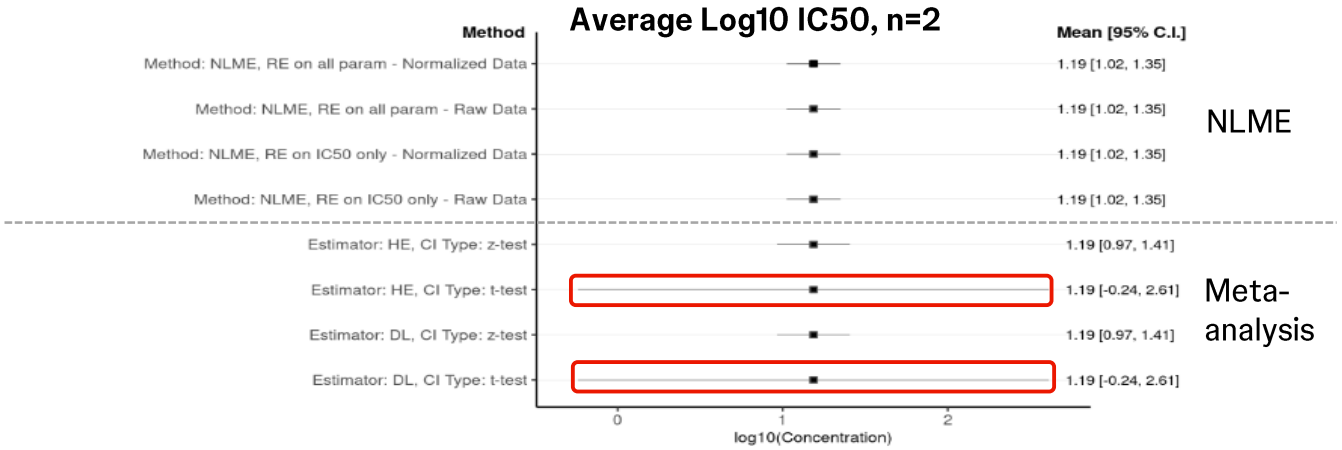


Simulation Results – Low Variation Between Curves

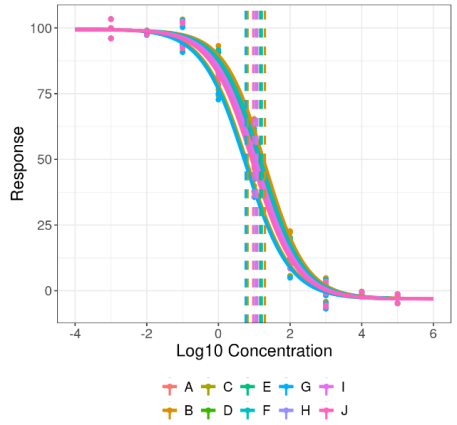
- Average IC50s nearly identical across methods for n=2 and n=10 experiments
- CIs similar for all methods except meta-analysis with t-test CIs for n=2

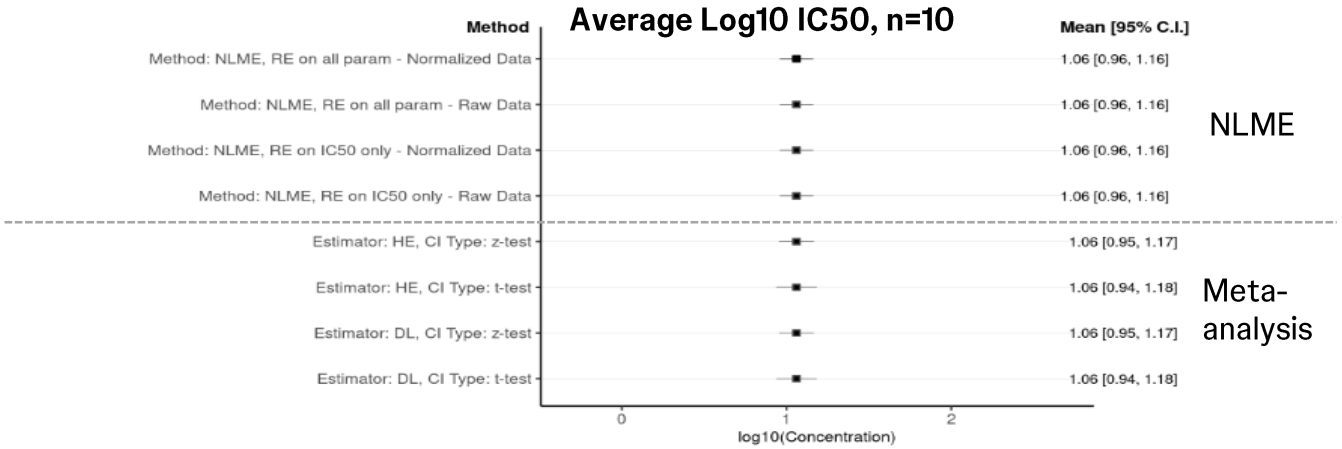
N=2 Experiments





N=10 Experiments

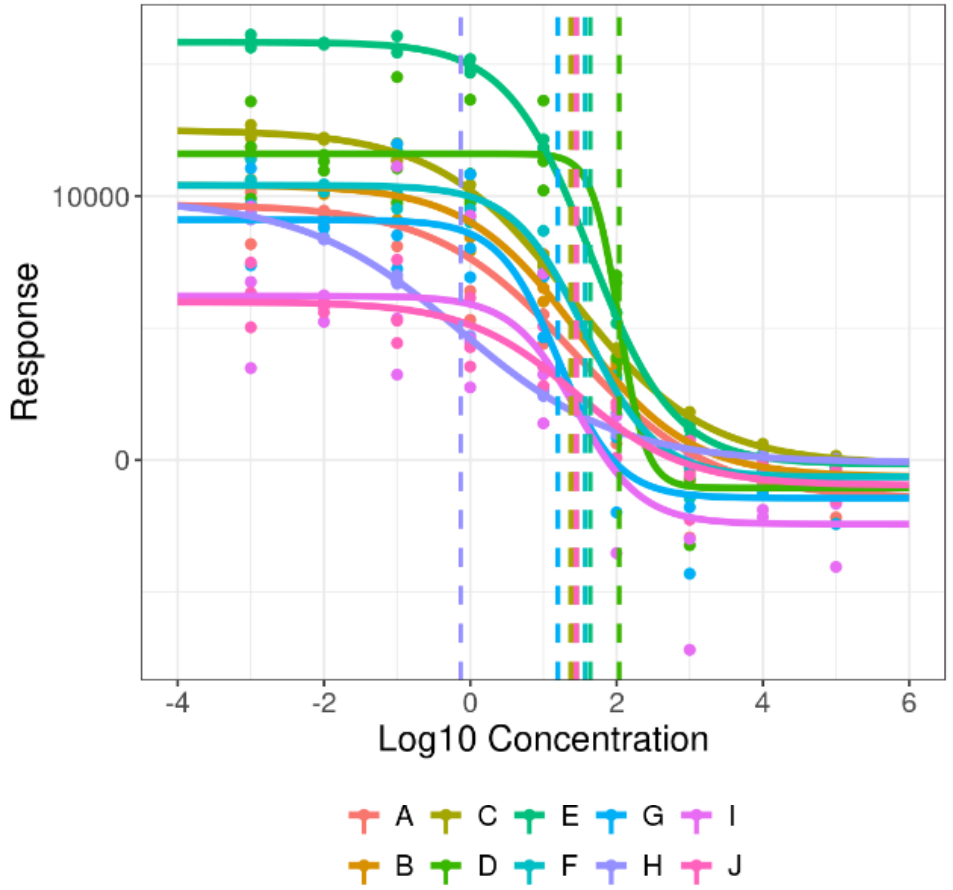




Examples Through Simulation – Different Scales

Simulation Setup – Curves on Different Scales

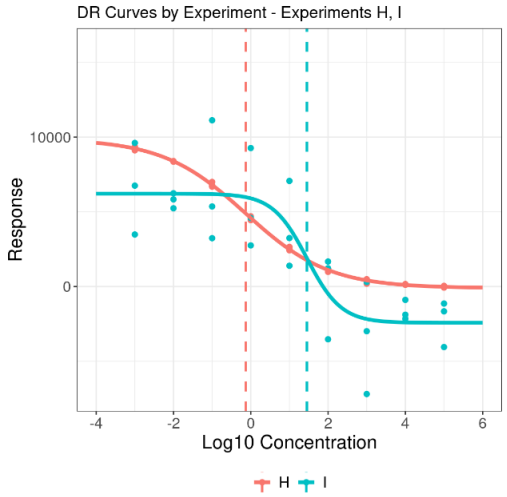
- Simulated dose-response curves for 10 experiments
- Higher variation between top parameters of curves
- Different within-experiment variation

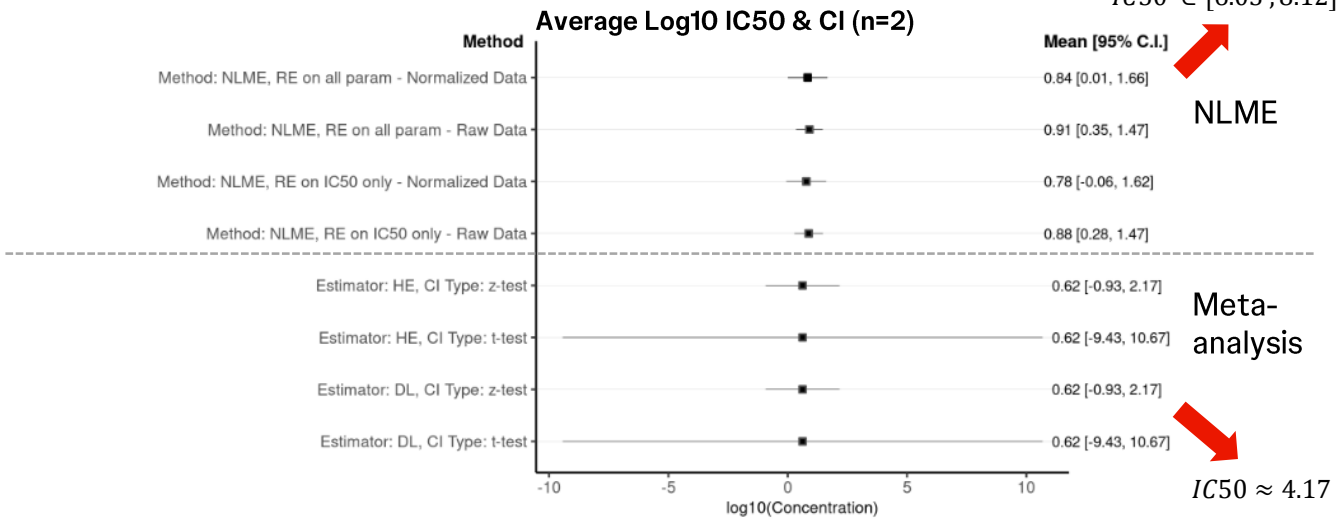


Simulation Results – Curves on Different Scales

- NLME on raw data w/o random effect on all parameters can produce very different average IC50 than other methods

N=2 Experiments





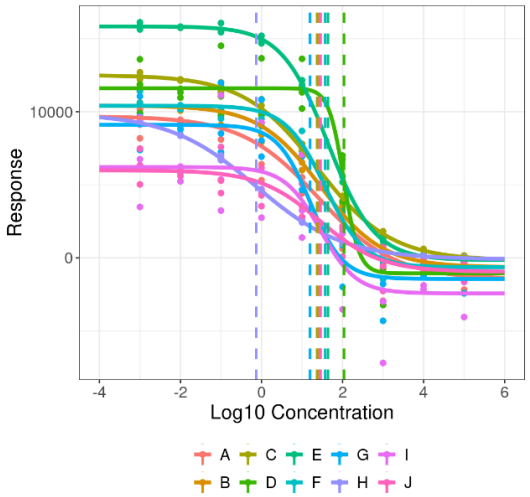
$IC50 \in [6.03, 8.12]$

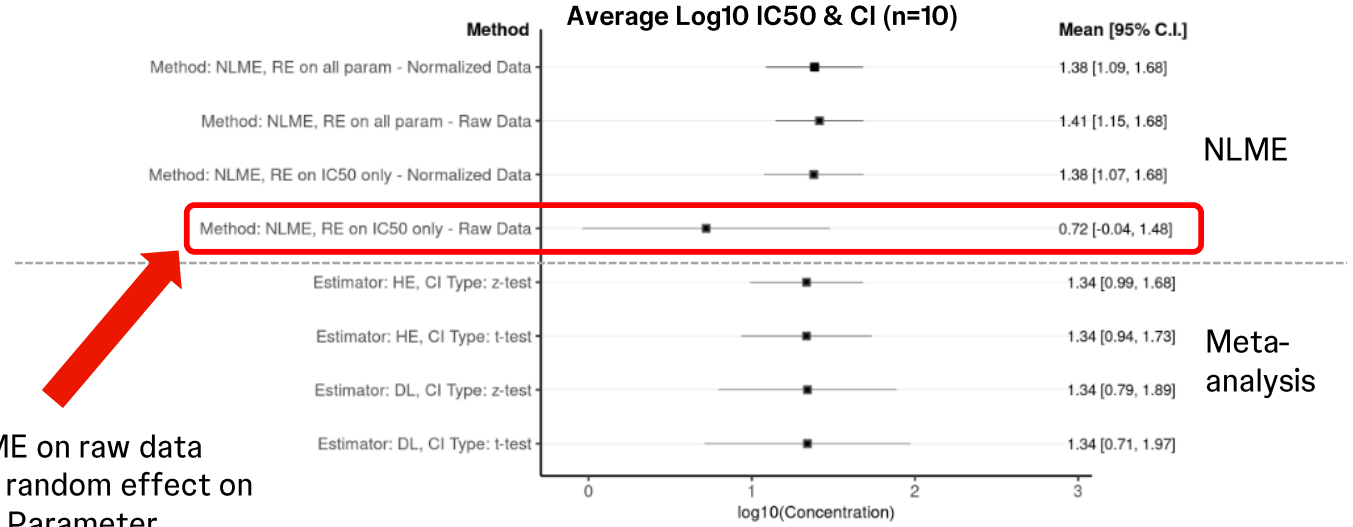
NLME

Meta-analysis

$IC50 \approx 4.17$

N=10 Experiments





NLME

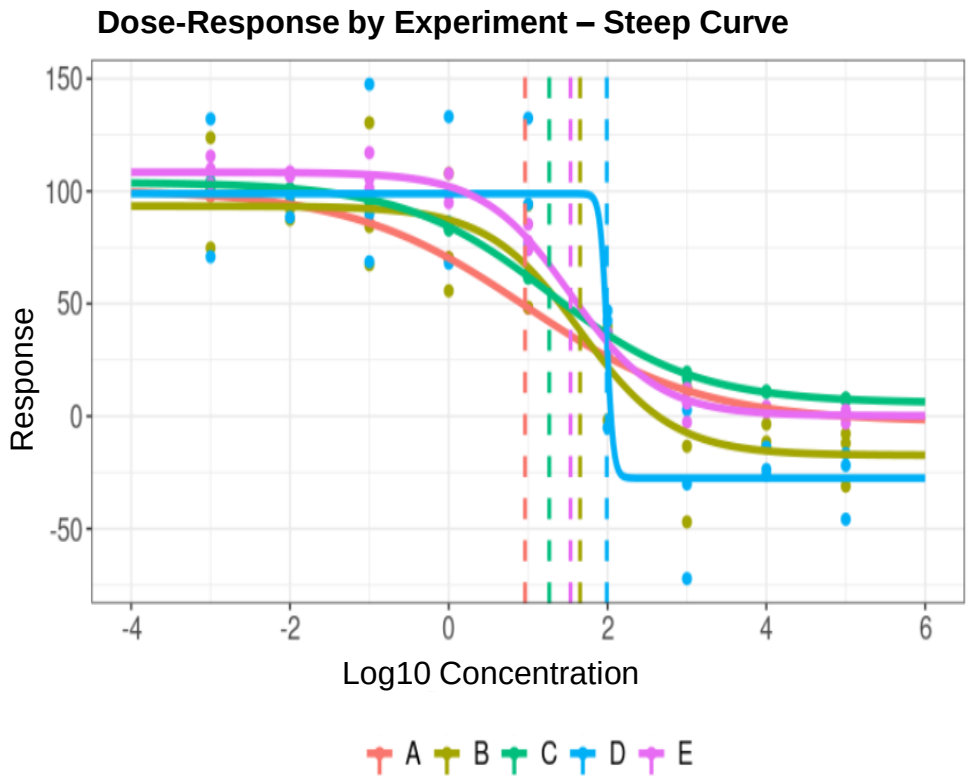
Meta-analysis

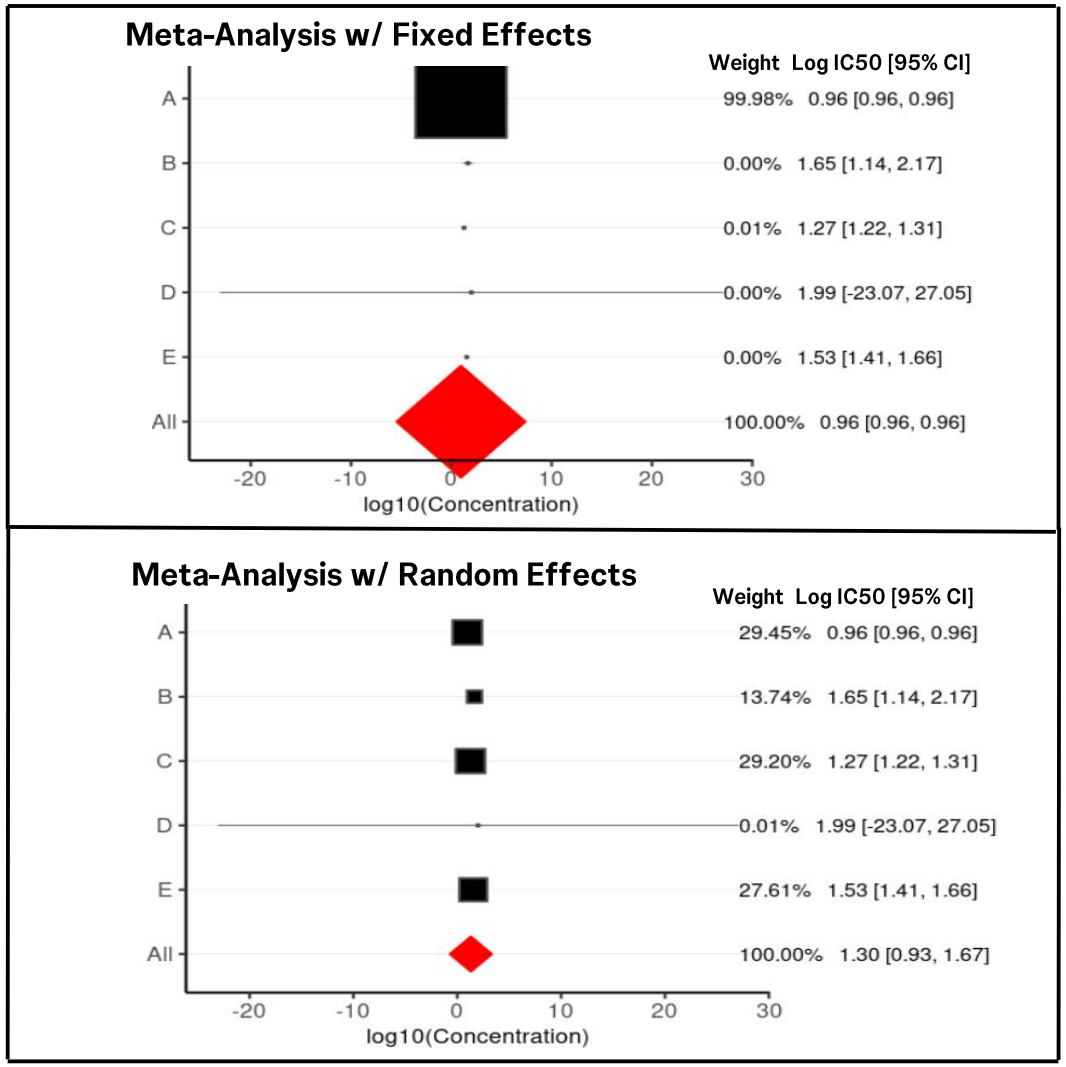
NLME on raw data w/o random effect on Top Parameter

Examples Through Simulation – Steep Curves

Simulation Setup – Steep Curves

- Steep curves wreck the weighting scheme in meta-analysis with fixed-effects
- Meta-analysis will remove experiments with steep curves, resulting in loss of information
- If many steep curves present, convergence issues may occur with NLME





Summary

Summary

- There is no ‘one size fits all’ approach
  - Discovery often has limited sample size and variable data
- Most of the methods produce similar average IC50 estimates
  - Except NLME on raw data without random effects on all parameters when curves are on different scales
  - Except for steep curves
- Meta-analysis with random effects is a reasonable approach under most circumstances

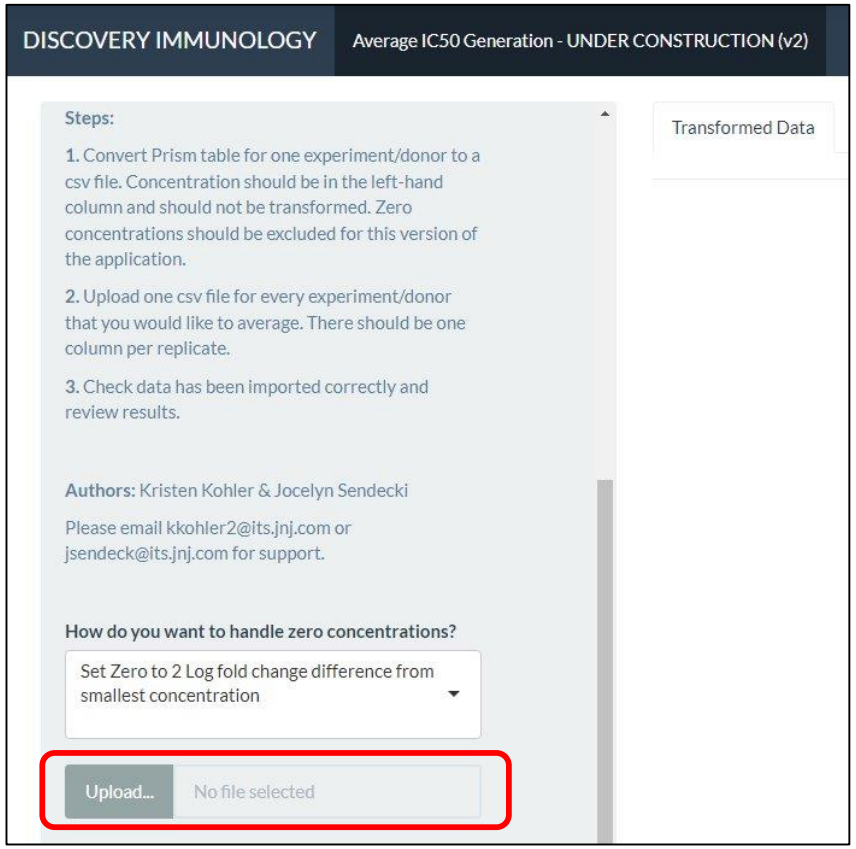
An Automated Solution

An Automated Solution

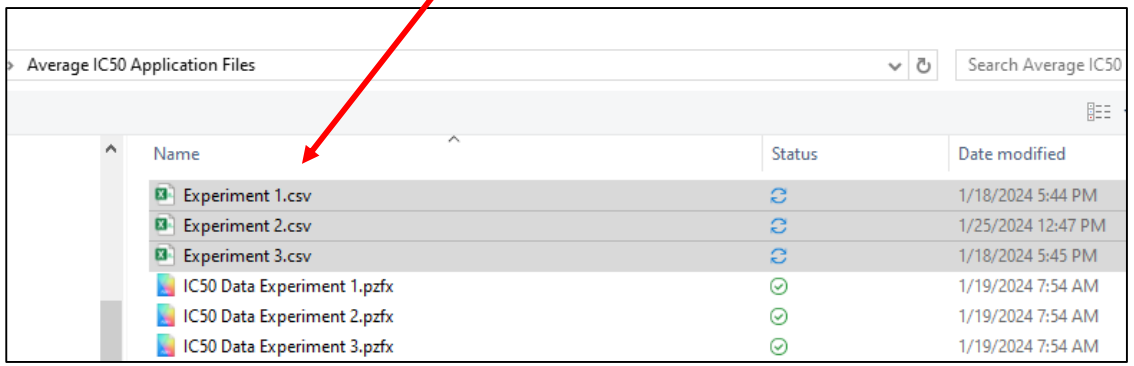
- Ultimately, we want an option to aggregate IC50s that is easy to understand and more statistically accurate
- Scientists may want to explore data on their own to get an approximate average IC50 across their experiments
- Creating application that will calculate IC50s by donor and average across donors using meta-analysis with random effects and z-test confidence intervals

Application Demo – Upload Data

- Upload csv file for each experiment



Choose all experiments

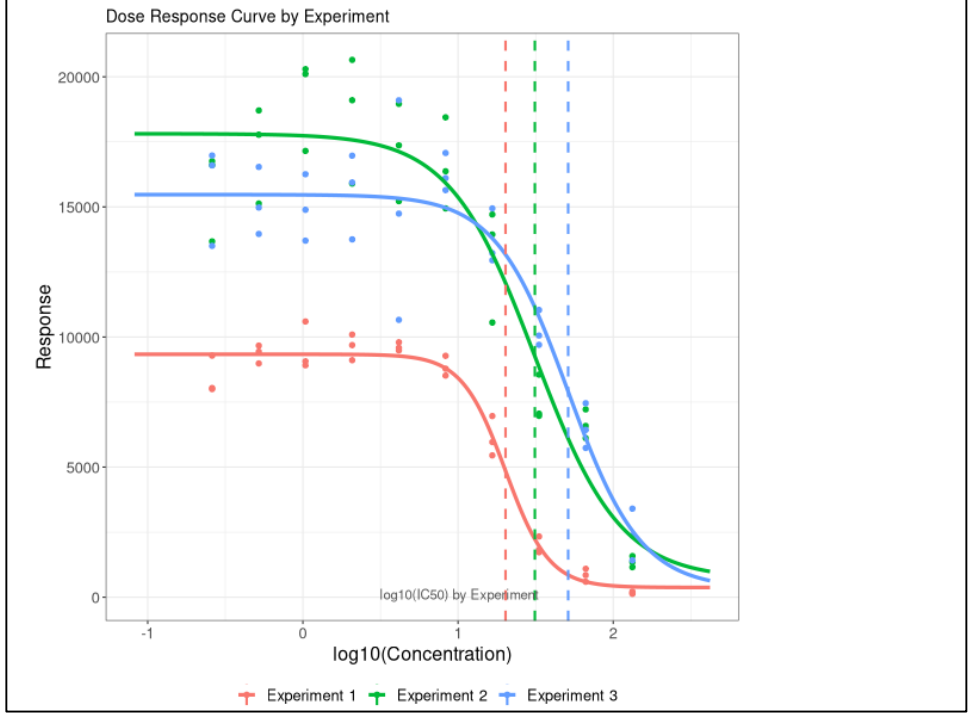


Application Demo – View IC50 by Experiment

- View IC50s and dose-response curves by experiment/donor

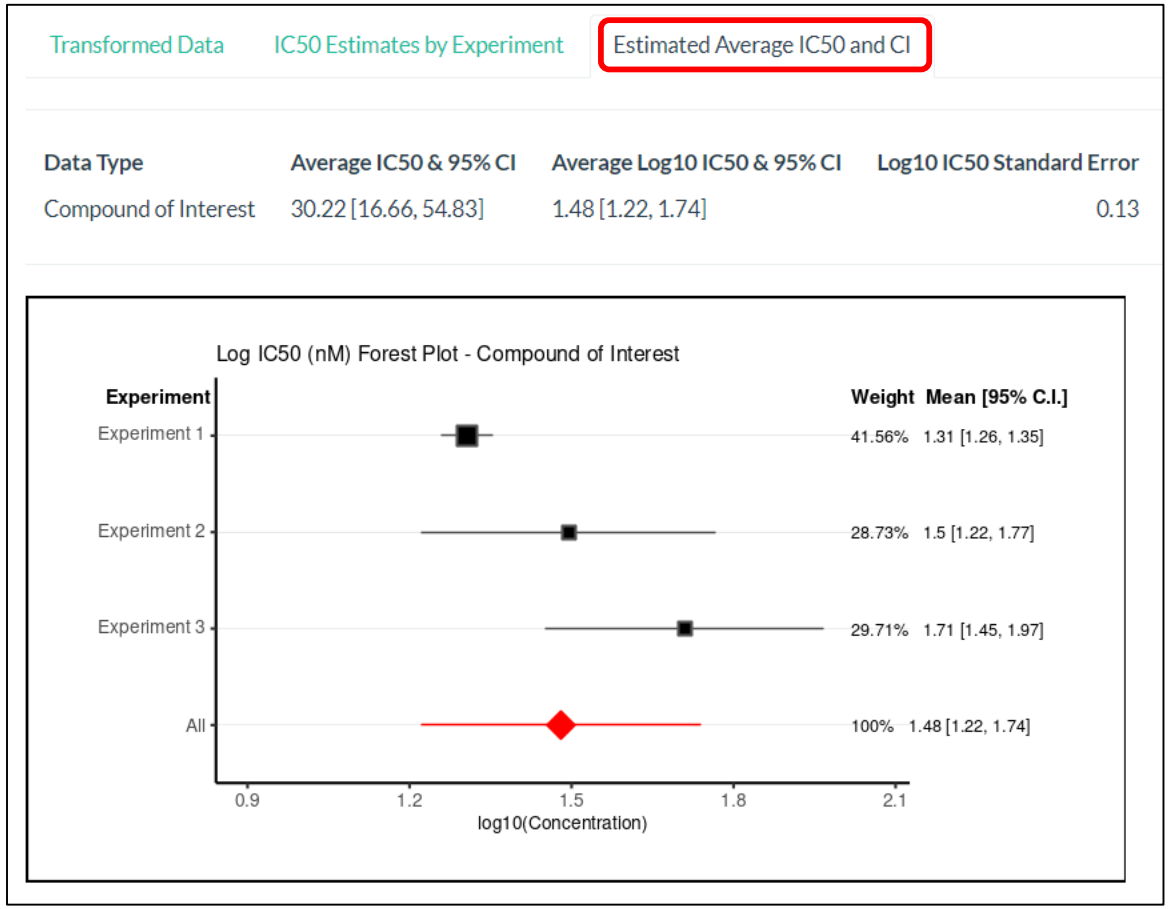
Transformed Data | IC50 Estimates by Experiment | Estimated Average IC50 and CI

Data Type	Experiment	IC50 & CI	Log10 IC50 & CI	Log10 IC50 Standard Error
Compound of Interest	Experiment 1	20.24 [18.1, 22.63]	1.31 [1.26, 1.35]	0.02
Compound of Interest	Experiment 2	31.26 [16.69, 58.57]	1.5 [1.22, 1.77]	0.14
Compound of Interest	Experiment 3	51.25 [28.29, 92.86]	1.71 [1.45, 1.97]	0.13



Application Demo – View Average IC50 & Forest Plot

- View weighted average IC50 calculated via meta-analysis and forest plot with results by experiment & overall



Future Application Add-ons

- Ability to compare IC50s of different treatments/compounds
- Ability to apply acceptance criteria
- Downloadable pdf report that could be uploaded to Electronic Lab Notebook (ELN)

Conclusion & Next Steps

Conclusion

- There is no ‘one size fits all’ approach
  - Discovery often has limited sample size and variable data
- Meta-analysis with random effects is a reasonable approach under most circumstances
  - Except for steep curves
- Shiny application can be used to help scientists get a more statistically accurate average IC50 estimate quickly

Next Steps

- Continue to explore methods and additional scenarios
  - Working group established in J&J Discovery Statistics to compare methods for special cases such as steep curves & non-responders
- Continue collaborating with scientists to build out application and deploy internally for use

Acknowledgements

Discovery and Nonclinical Safety Statistics

- Jocelyn Sendecki
- Traymon Beavers
- Nicholas Hein
- Bie Verbist
- Jeroen Tolboom
- Fetene Tekle
- Yannick-Andre Breton

Discovery Immunology

- Paul Dudas
- John DeLong
- Clara Moon
- Kacey Sachen
- Indra Sarabia
- Astrid Clarke
- Janise Deming

Thank you

If you have any questions, please contact:

Kristen Kohler
kkohler2@its.jnj.com

References

- *'Summarizing EC50 Estimates from multiple dose-response experiments: A comparison of a meta-analysis strategy to a mixed-effects model approach'*, Jiang X. , Kopp-Schneider A.

Back-Up Slides

Code Snippets – Curve Simulation Example

```
experiment_num = 10

var.vec <- c(1500, 700, 200, 1500, 200, 700, 1500, 100, 2500, 1000)
data_sim_diff_var = NULL

for (k in 1:experiment_num){

  set.seed(k*100)

  x_vec = -3:5
  A_sim1 = 15 + rnorm(1,0,5)
  B_sim1 = 10000 + rnorm(1,0,3000)
  lic50_sim1 = 1 + rnorm(1,0,.7)
  slope_sim1 = -.75 + rnorm(1,0,.4)

  rep_num = 3

  y_sim1 = NULL

  for (i in 1:rep_num){

    set.seed(experiment_num*i)

    var.tmp <- var.vec[k]

    y_sim1 =
      c(y_sim1,
        A_sim1 + (B_sim1-A_sim1)/(1+10^(slope_sim1*(lic50_sim1 - x_vec))) +
          rnorm(length(x_vec),
                0,
                var.tmp))

  }
```

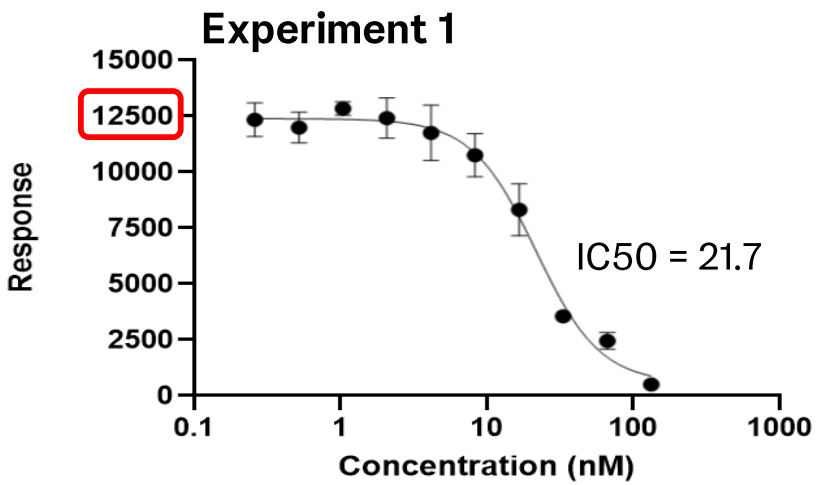
Code Snippets – NLME Example

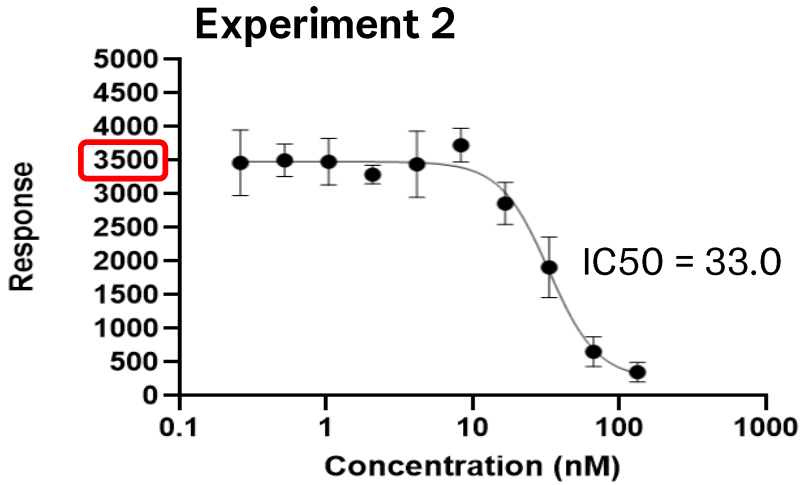
```
#set convergence criteria
control = nlmeControl(maxIter = 100,    msMaxIter = 1000,    tolerance = 0.1,   abs.tol = 1e-20)
```

```
eval(parse(text=paste0("analysis.data <- data.modeling %>% filter(Experiment.num <= ", i+1 , ")" )))

 TOLVAL = 0.0001
  YVAR = "Response"
  XVAR = "Log10.Concentration"

  currentFormula <- as.formula(paste(YVAR,"~ SSfpl(",XVAR,", A, B, xmid, scal)"))

  forAnalysis <- analysis.data

  currentNls <- try(nls(currentFormula, data = forAnalysis),silent=TRUE) # this gets your nlme starting values
  startVals <- coef(currentNls)
  currentModel <- try(nlme(currentFormula,
                           data = forAnalysis,
                           groups = ~Experiment,
                           fixed = A + B + xmid + scal ~ 1,
                           control = control,
                           random = pdDiag(xmid ~ 1),
                           # random = pdDiag(A + B + xmid + scal ~ 1),
                           start = list(fixed = startVals)),
                      silent = TRUE)
```

Practical Example with Method Comparison

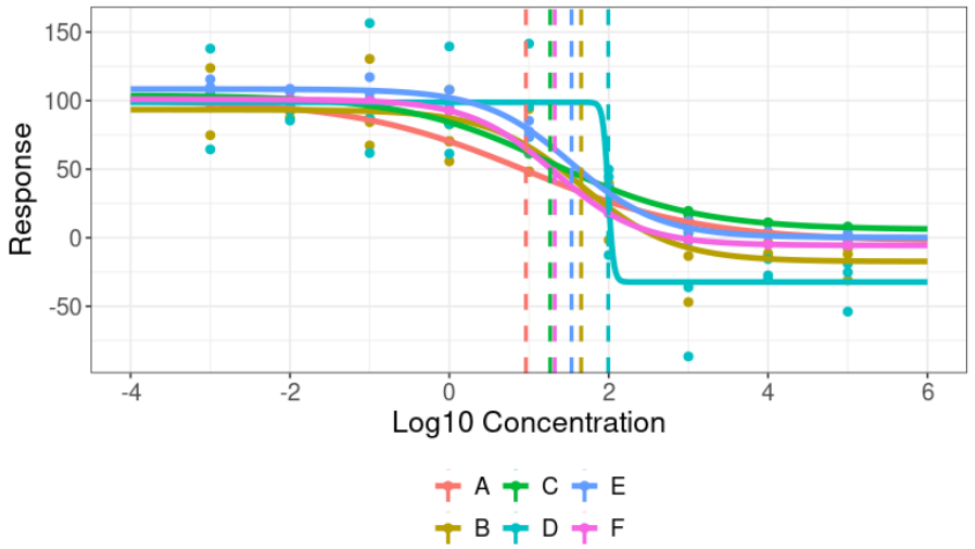
Example 2: Inhibition Data for Compound A (Species 2) – Different Scales

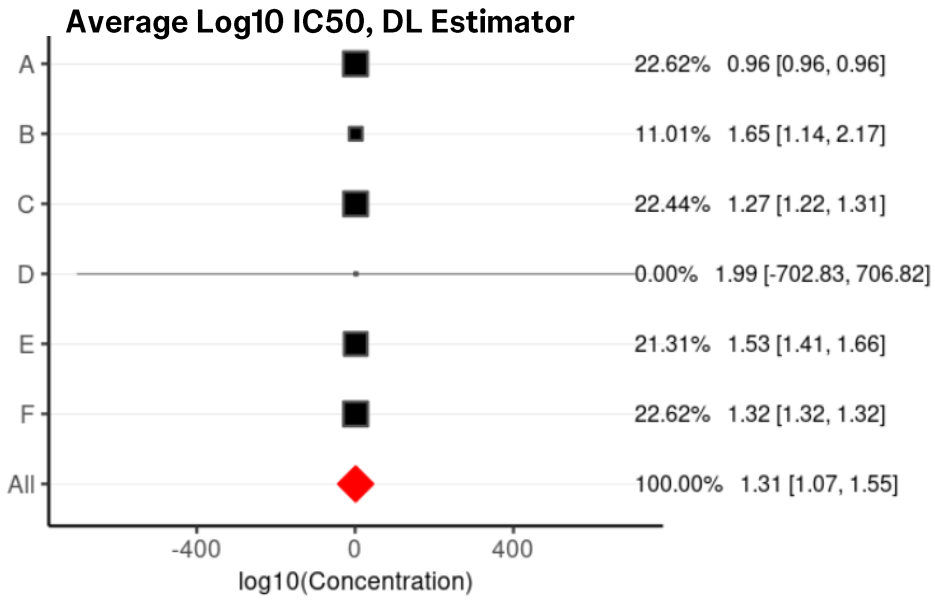


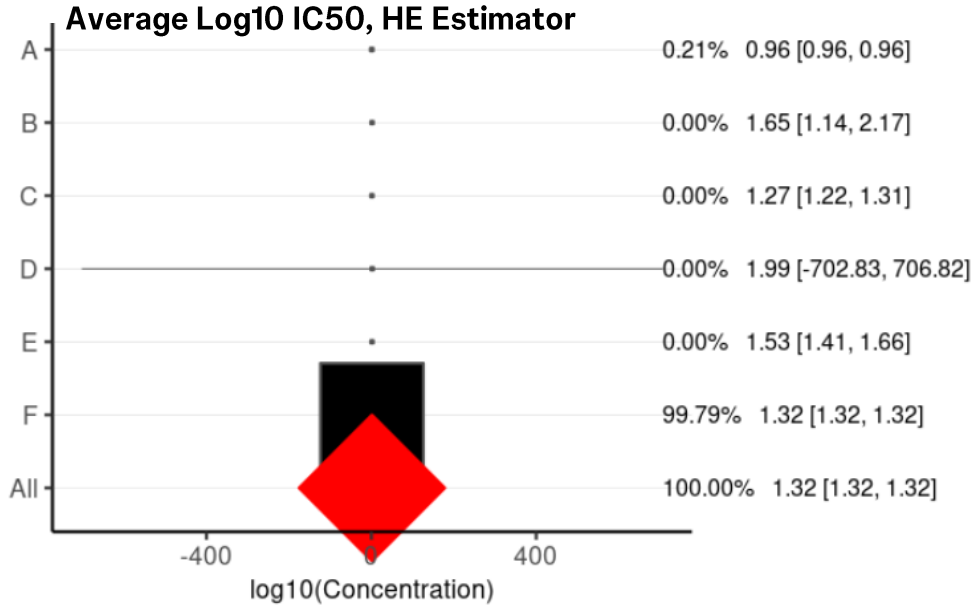


Method	Avg Log IC50 & CI
Meta-analysis w/ **random** effect (RE) & **t-test** CI	1.42 [.26, 2.59]
Meta-analysis w/ **random** effect (RE) & **z-test** CI	1.42 [1.24, 1.60]
Meta-analysis w/ **fixed** effects & **t-test** CI	1.42 [0.26, 2.59]
Meta-analysis w/ **fixed** effects & **z-test** CI	1.42 [1.25, 1.60]
NLME on **raw** data, **RE on IC50 only**	**0.19 [-1.32, 1.71]**
NLME on **raw**, data **RE on all parameters**	1.35 [1.30, 1.34]
NLME on **normalized** data, RE on IC50 only	1.42 [1.31, 1.53]
NLME on **normalized** data, RE on **all parameters**	1.42 [1.29, 1.55]

Curves with Convergence Issues







Method	Avg Log IC50 & CI
Meta-analysis w/ random effect (RE)	1.31 [1.07, 1.55]
Meta-analysis w/ fixed effects	1.32 [1.32, 1.32]
NLME on raw data, RE on IC50 only	1.56 [1.35, 1.77]
NLME on raw, data RE on all parameters	Does Not Converge
NLME on normalized data, RE on IC50 only	1.48 [1.27, 1.70]
NLME on normalized data, RE on all parameters	1.52 [1.29, 1.74]

Meta-analysis Equations

Weighted Average

$$\hat{\mu} = \sum_{i=1}^{k} \hat{w}_i \phi_i^{(e)} \Big/ \sum_{i=1}^{k} \hat{w}_i$$

Weights

$$w_i = 1/(\tau^2 + \sigma_i^2)$$

Estimators

Hedges' Estimator (fixed effects)

$$\hat{\tau}^2_{HE} = max\left\{0, \frac{1}{k-1}\sum_{i=1}^{k}\left(\phi_i^{(e)} - \frac{1}{k}\sum_{i=1}^{k}\phi_i^{(e)}\right)^2 - \frac{1}{k}\sum_{i=1}^{k}\hat{\sigma}_i^2\right\}$$

$$\hat{\tau}^2_{DL} = max\left\{0, \frac{\sum_{i=1}^{k} m_i(\phi_i^{(e)} - \overline{\phi^{(e)}}_m)^2 - (k-1)}{\sum_{i=1}^{k} m_i - \sum_{i=1}^{k} m_i^2 / \sum_{i=1}^{k} m_i}\right\}$$

DerSimonian-Laird (DL) Estimator (random effects)