Limiting potency bias from allowed non-similarity while protecting the similarity pass rate

David Lansky, Ph.D., david@precisionbioassay.com

precision bioassay SM inc

Burlington, Vermont, USA

October 2024

Abstract

© precision bioassay inc

D. Lansky

Protecting bioassay-based estimates of relative potency against bias (due to allowed non-similarity) and against unacceptably high similarity failure rates, while allowing for changes in assay capability (precision) appears to be impractical. While power calculations for both detecting non-similarity (via difference tests) and for passing similar samples (via equivalence tests) are particularly helpful, these calculations are not simple. This presentation will illustrate some tools for these calculations and how they support modifying similarity criteria based on data from bioassay.

Similarity is essential in bioassay

© precision bioassay inc

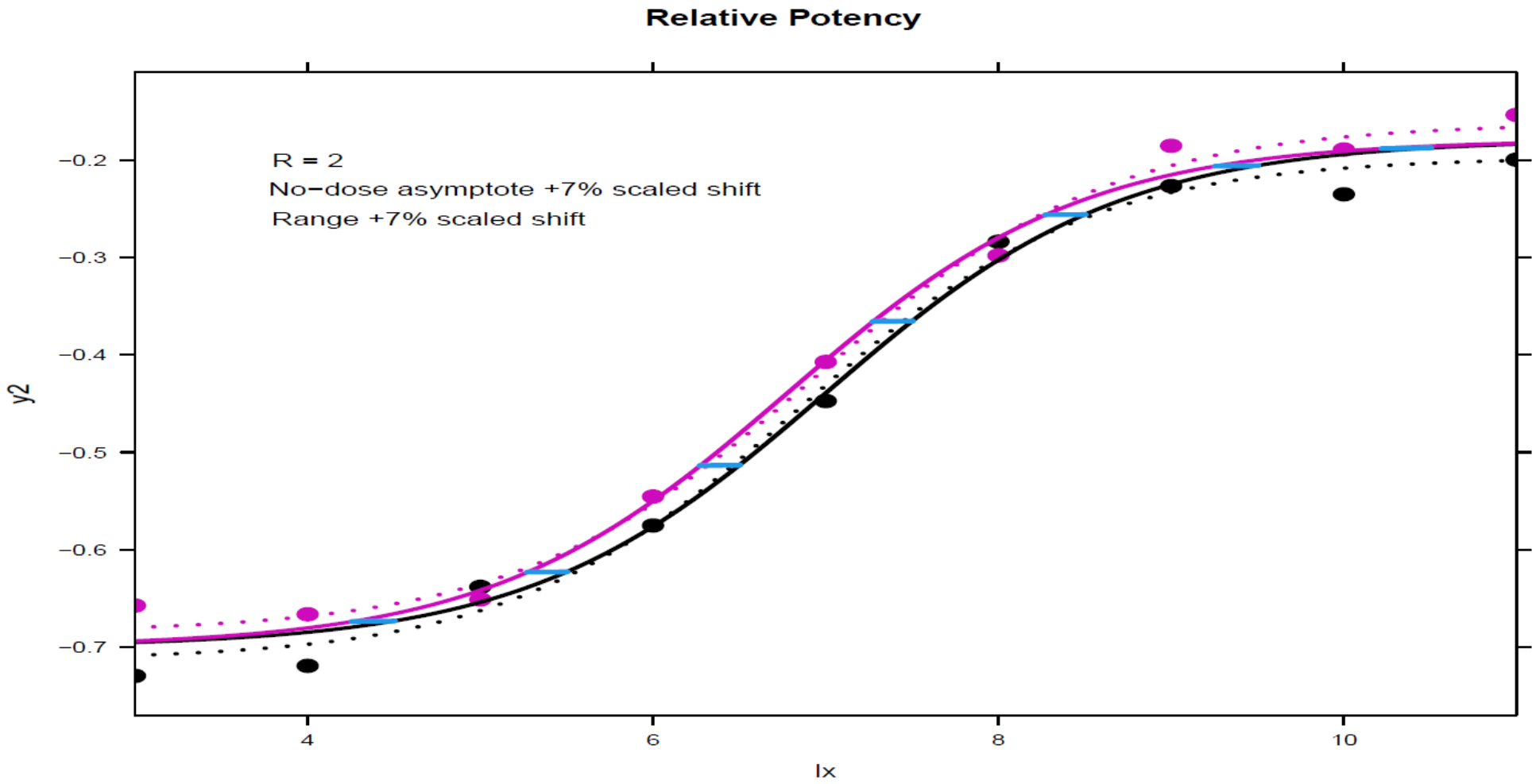
D. Lansky

- Potency isn't meaningful without biological similarity (= samples contain same [mixture of] active compound[s])
- Biological non-similarity is important SAMPLE quality information
- Assay (not sample) non-similarity is important ASSAY quality information
- Statistical similarity is necessary, but not sufficient
- Small amounts of some types of non-similarity cause important potency bias

Unconstrained vs. Constrained Model

© precisionbioassay inc

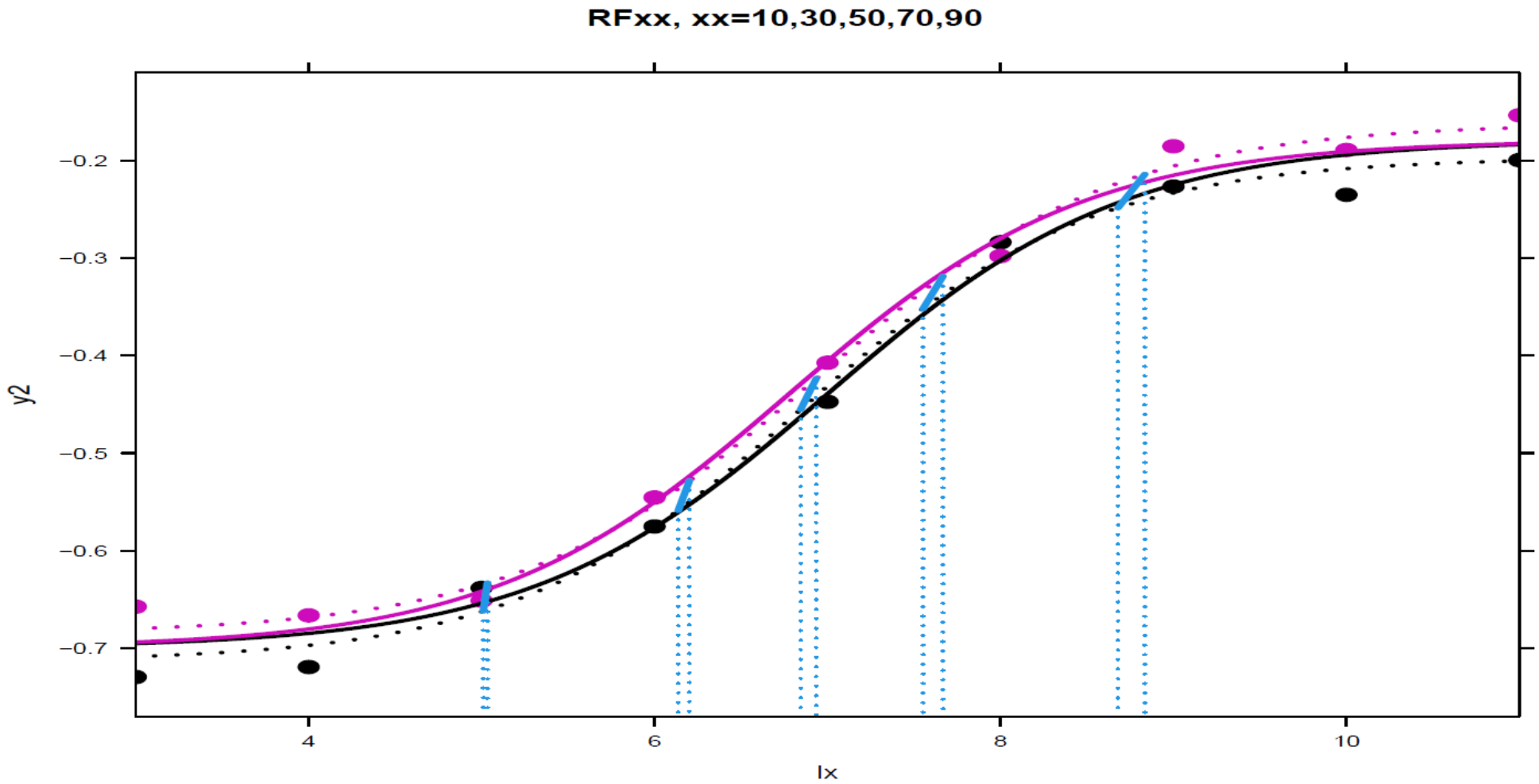
D. Lansky



Black = ref, magenta = test, dots = unconstrained, solid = constrained

RFxx = EDxx comparisons w/unconstrained model

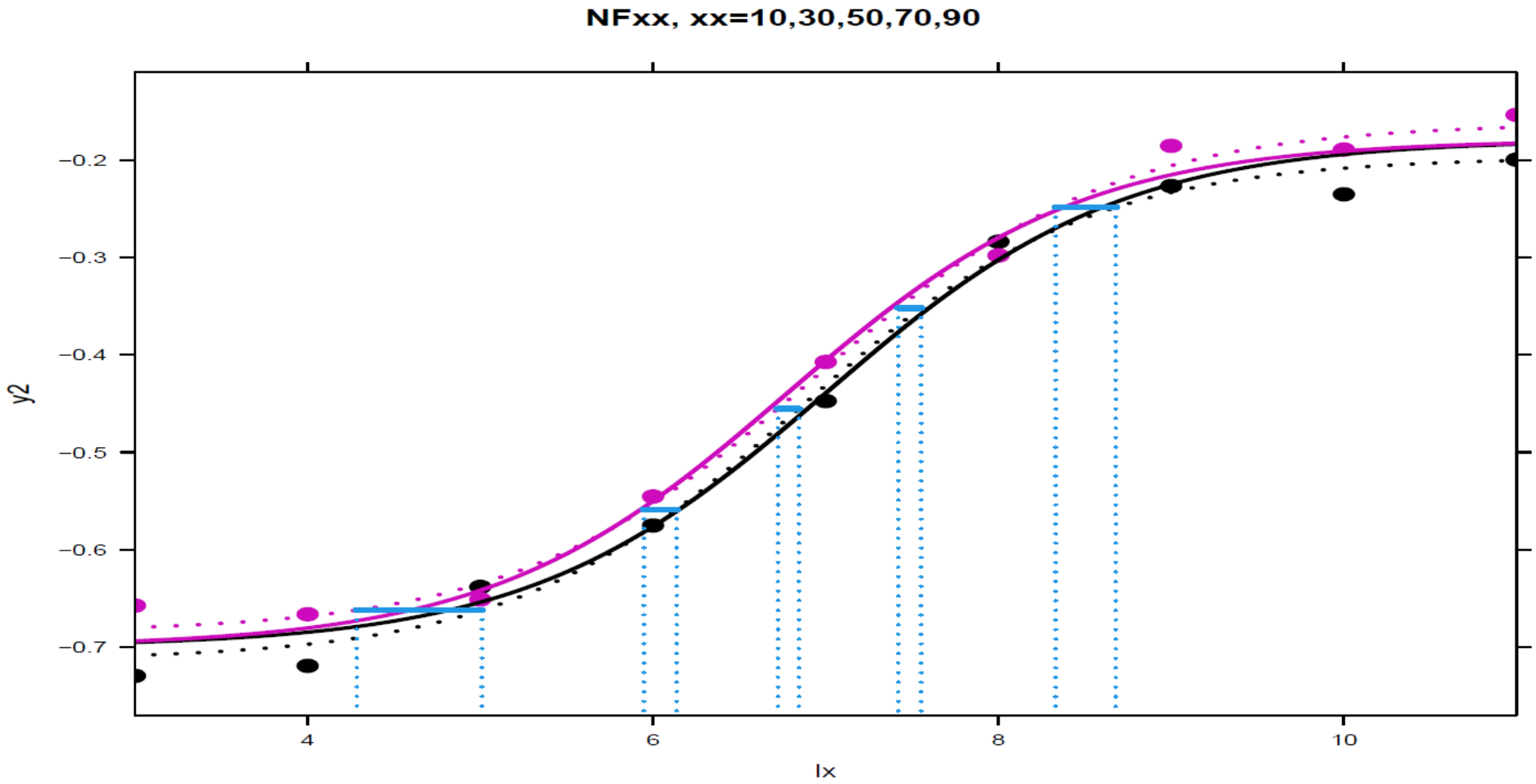
D. Lansky



Black = ref, magenta = test, dots = unconstrained, solid = constrained

NFxx compares doses at response of ref EDxx

D. Lansky



Black = ref, magenta = test, dots = unconstrained, solid = constrained

Relative Potency vs. RFxx and NFxx

© precisionbioassay inc

D. Lansky

- Relative potency uses constrained model (assumes biological similarity of samples and no assay-induced non-similarity)
- RFxx simply compares EDxxs from unconstrained model (typically useless)
- NFxx compares doses at specified response levels of ref using unconstrained models (useful in special cases)
- Both RFxx and NFxx vary with xx; what xx to use and why?
- While RFxx or (more likely) NFxx may be useful measures of activity, neither estimates potency

With biological similarity relative potency robust

© precision bioassay inc

D. Lansky

- With biological similarity:
  - Expect different bioassay systems (that respond to target) to deliver the same relative potencies
  - Surprises (samples thought to be biologically similar that delivered different relative potencies in different bioassay systems) led to discovery of different subgroups of some antibiotics (multiple bioassays useful in discovery/development?)
- RFxx and NFxx are fragile (very sensitive to):
  - differences in the (mixtures of) active compound(s),
  - dose level,
  - response level, and
  - the assay system

Scaled Shift Non-Similarity

© precisionbioassay inc

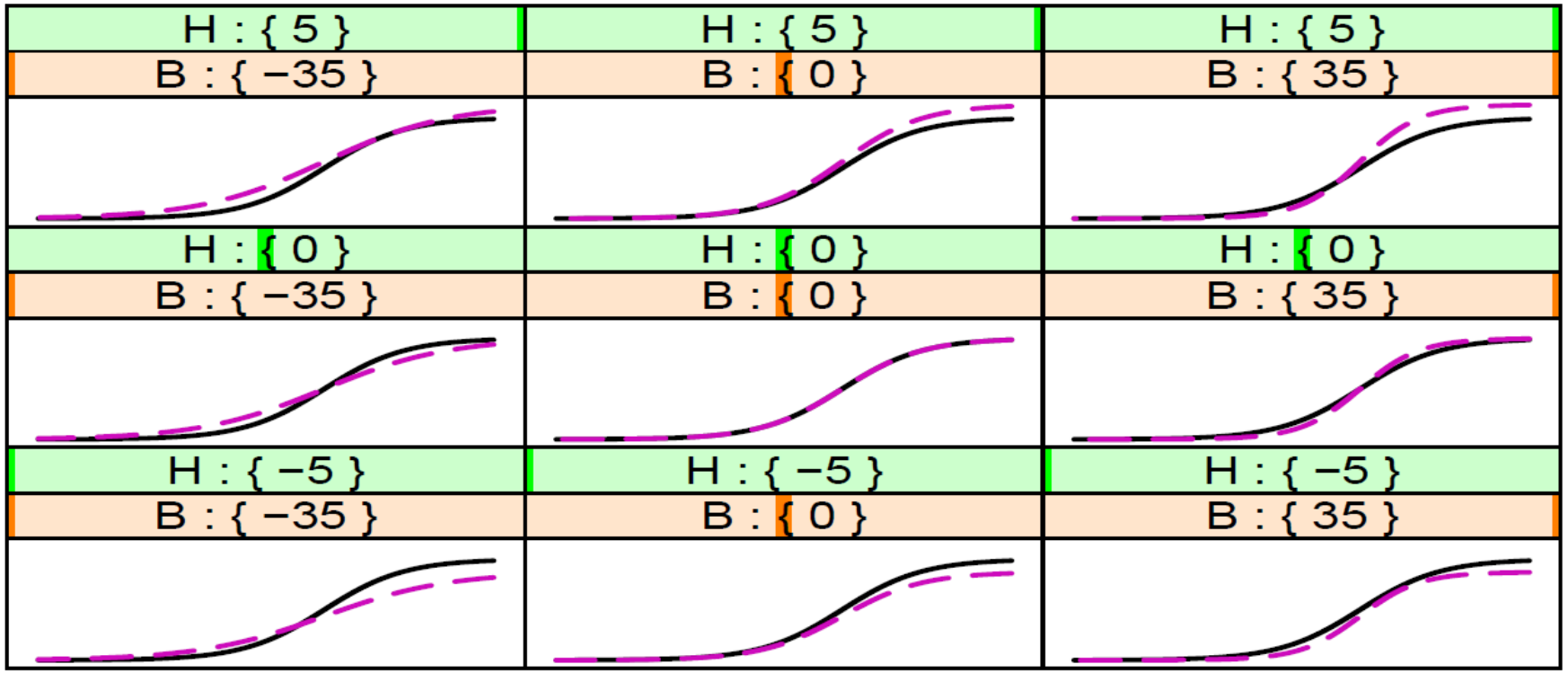
D. Lansky

- $y^* = \frac{H_i}{1+e^{-B_i(\ln(x)-C_i)}} + D_i + \epsilon$
  - H Response Range
  - B "Slope"
  - C Ln EC50
  - D No-dose Asymptote
  - Note: $H = A - D$ where A is max response asymptote
- Responses & parameter units vary across assays
- Sometimes useful to use $\log_2$ rather than ln
- Scaled Similarity Parameters
  - $\%\Delta_H = 100 \times (H_{\text{Test}} - H_{\text{Ref}}) / \overline{H^*_{\text{Ref}}}$
  - $\%\Delta_D = 100 \times (D_{\text{Test}} - D_{\text{Ref}}) / \overline{H^*_{\text{Ref}}}$
  - $\%\Delta_B = 100 \times (B_{\text{Test}} - B_{\text{Ref}}) / \overline{B^*_{\text{Ref}}}$

* Long term average

No-dose Asymptote 0% shift

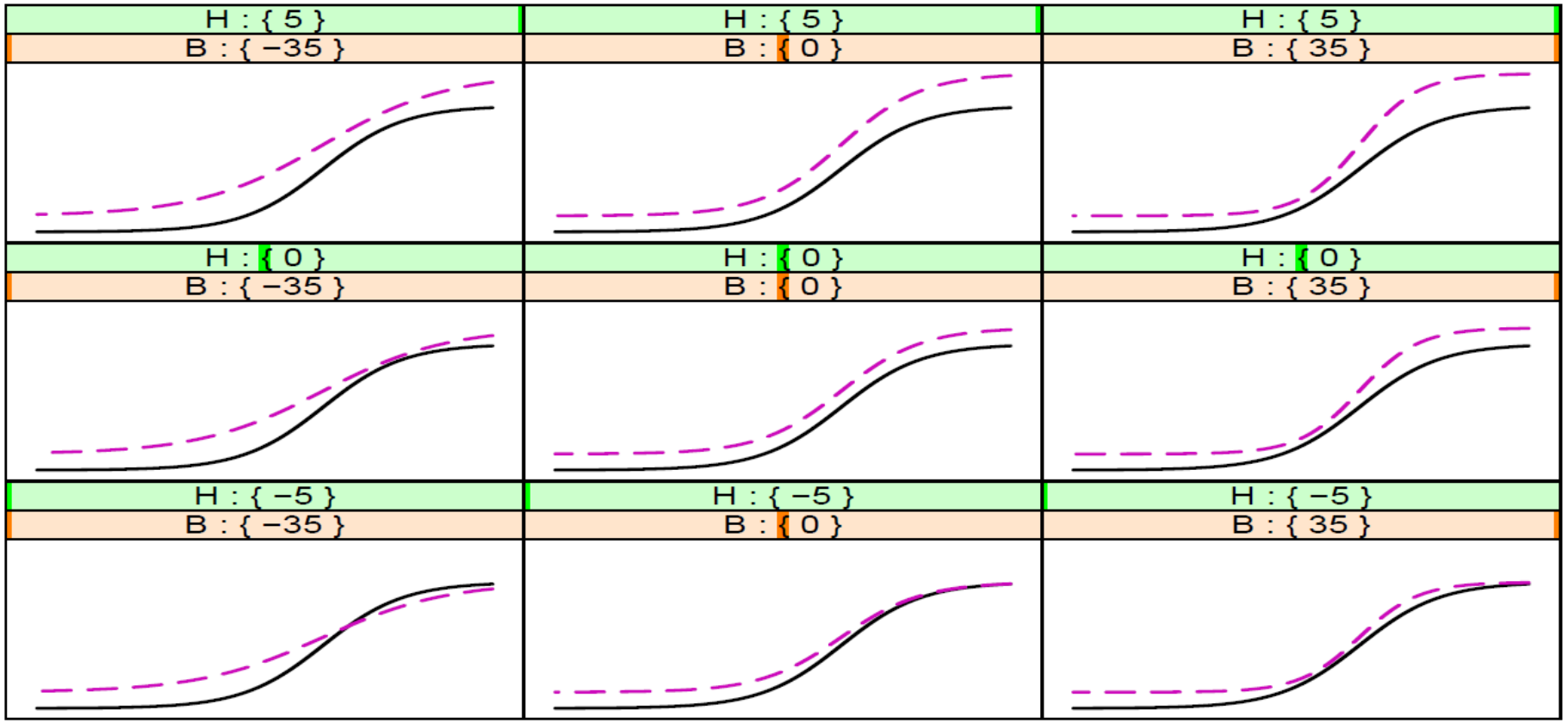
D. Lansky



Black is reference, magenta is test, all curves have potency 1

No-dose Asymptote +5% shift

D. Lansky



Black is reference, magenta is test, all curves have potency 1

"Scaled shift" equivalence bounds: Experience

© precisionbioassay inc

D. Lansky

- $\%\Delta_{H:B:D:A}$ range:effect rate:no-dose asymptote:max response asymptote
  - Even excellent assays struggle with 5:35:5:5
  - Noisy assays struggle with 10:50:10
  - 10:50:10:10 almost keeps median bias under 10%
  - linking potency bias to equivalence bound - challenging
- Challenging to choose equivalence bounds:
  - control potency bias and
  - have adequate power to pass similarity
- Solution:
  - Use potency bias limits to set **difference** test bounds for nonsimilarity
  - Use stat properties of diff & equiv tests (power) to set equiv bounds
  - Use sound monitoring and adapt as needed

Simulated Bioassays

© precisionbioassay inc

D. Lansky

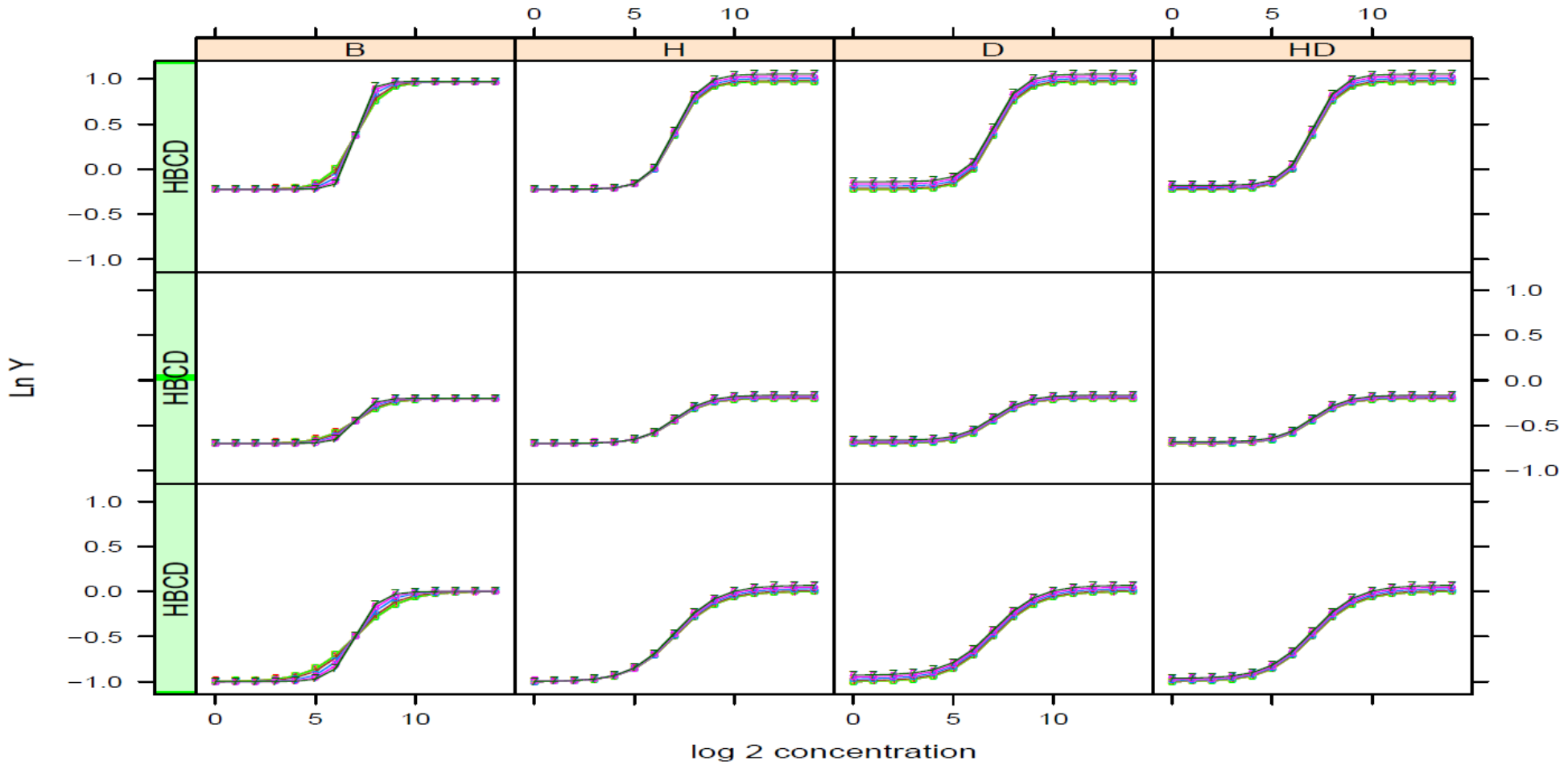
- 3 sets of curve parameters:

	H	B	C	D
HBCD3	1.19	1.50	7.00	-0.22
HBCD2	0.50	1.20	7.00	-0.70
HBCD1	1.00	0.90	7.00	-1.00

- 5 levels of residual SD (as % of H): $2, 3, 4, 5, 6$
- Non-sim: Range (H), effect (B), no-dose asymptote (D), and HD (=A)
- Amount of non-similarity (as percent scaled shifts):
  - for D and H: $0, 1, 3, 5, 7$
  - for B: $0, 7, 21, 35, 49$
- Test samples all have relative potency of 1
- 16 replicate assays for each combination of conditions
- Fit unconstrained & constrained 4 PL to each assay, estimate scaled shift non-similarity and (assuming similarity) estimate potency

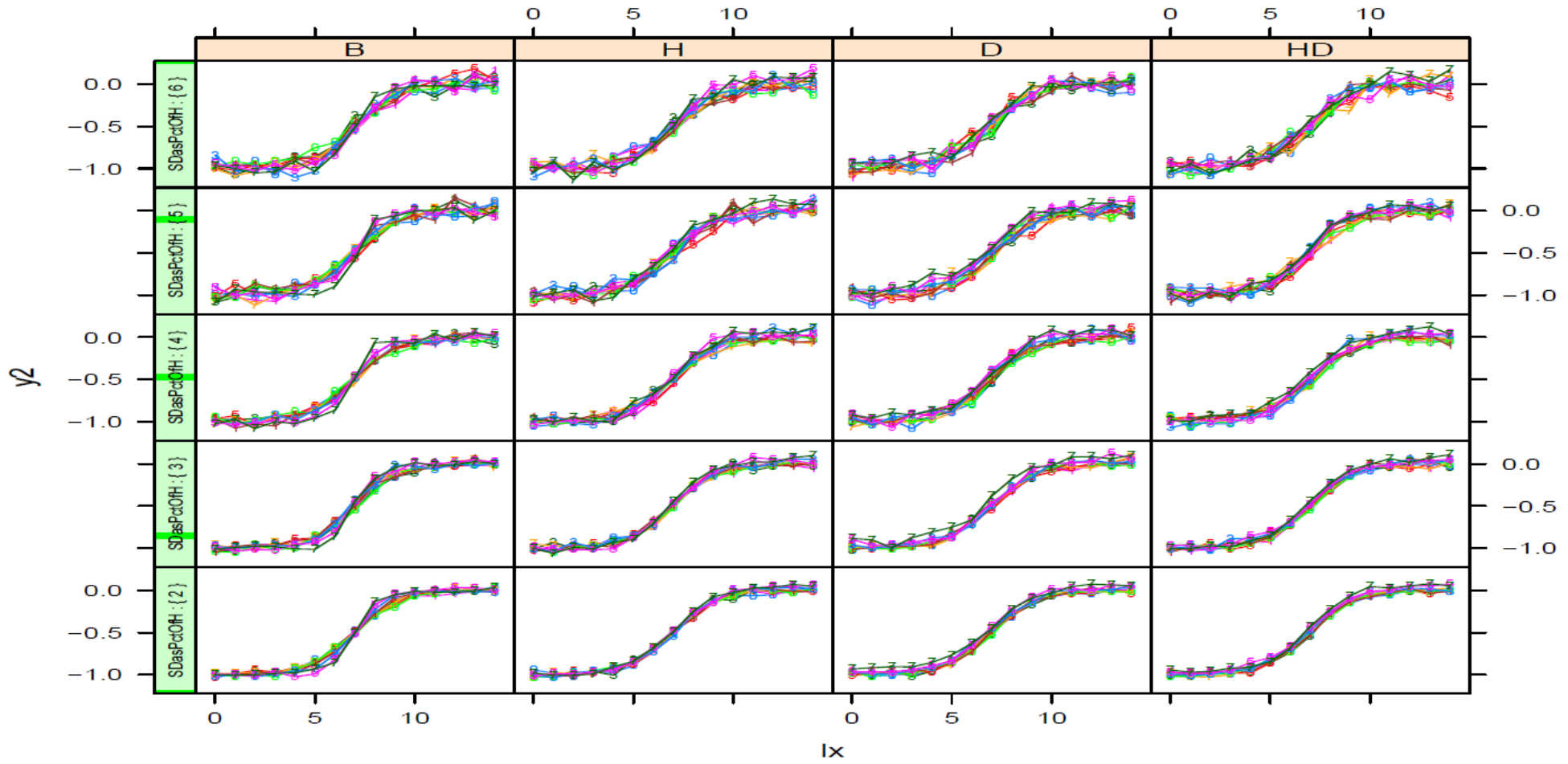
Impact of curve parameters and non-similarity

D. Lansky



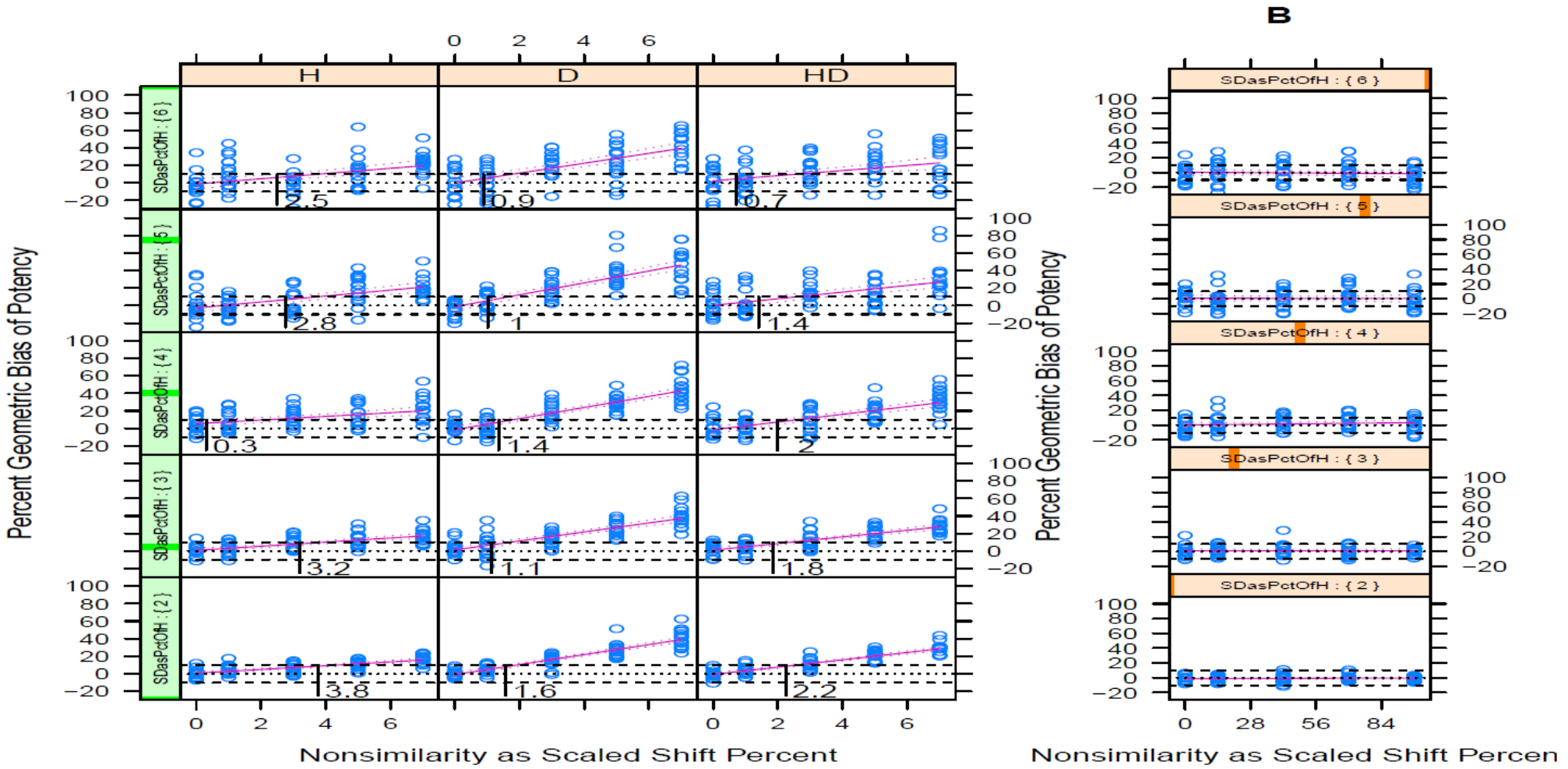
Impact of non-similarity and SD: HBCD1 (bottom row)

D. Lansky



Impact of SD and non-similarity on Potency

D. Lansky

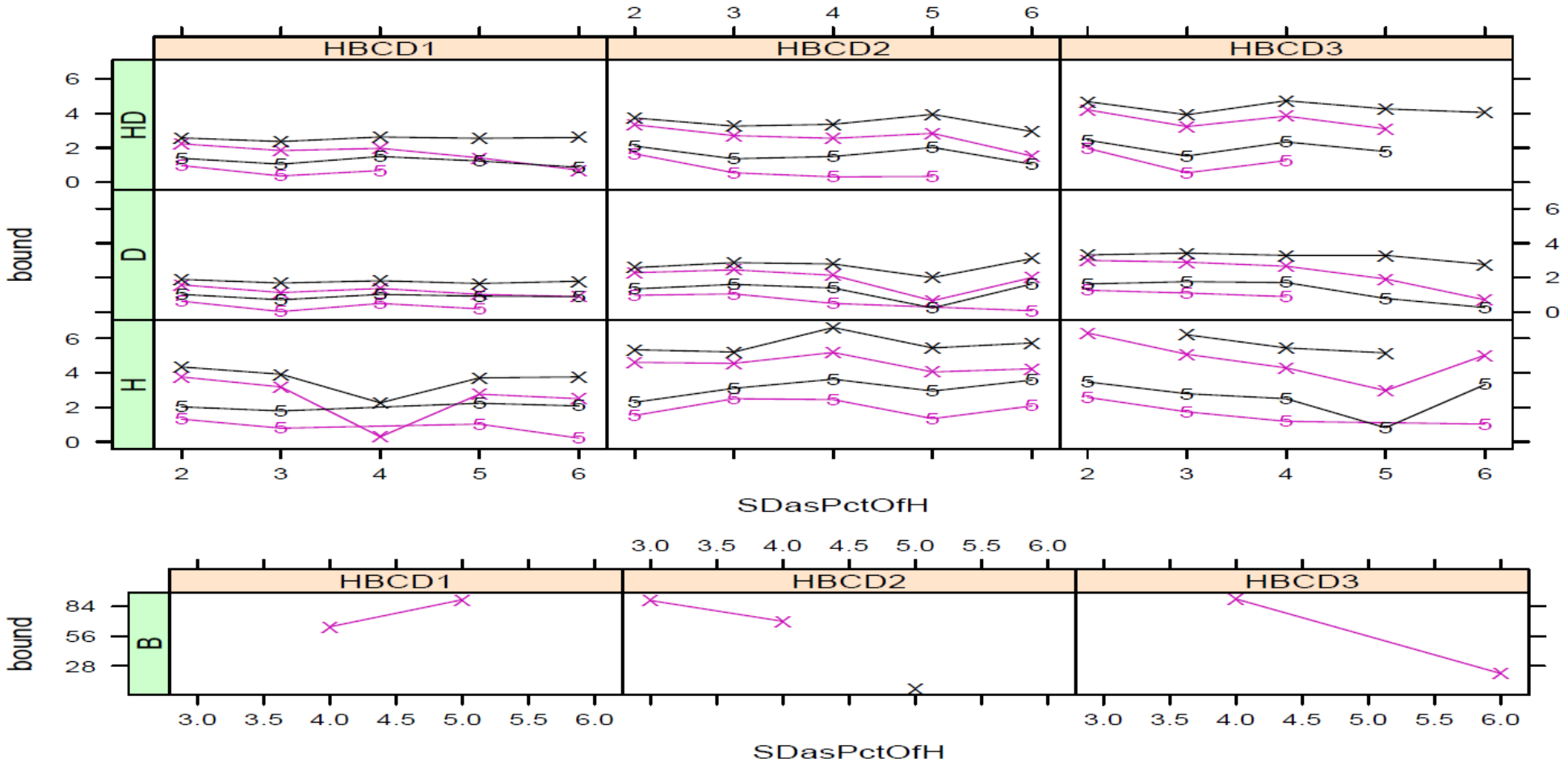


Extract x from above for mean and upper 95% CI

- Different combinations of curve shape (slide 13) and
- Residual SD (as % of range)
- For each non-similarity parameter
- Allowing 5% (5) or 10% (X) geometric bias in estimate relative potency
- For the mean (black) or upper 95% CI of mean (magenta)

Difference test bounds to prevent potency bias

D. Lansky



Difference test bounds for scaled shift non-similarity

© precisionbioassay inc

D. Lansky

- Bounds have little sensitivity to curve shape or residual SD
- With residual SD $\leq$ 4% of response range

Table: Difference test bounds for mean scaled shift non-similarity by limit on percent geometric bias of potency and non-similarity measure

	B	D	H	HD
10	60.00	1.70	3.90	2.40
5	60.00	0.20	1.80	1.10

Bioassay: Big Picture

D. Lansky

- Essential to control potency bias with: design, development, method (routine randomization), analysis, monitoring
- Small amounts of (some types of) non-similarity cause bias in potency
- Use assay capability (σ of non-similarity estimates) and product knowledge or bias limit to set assay size (n)
- Consider shifting similarity assessment from method (assay) to procedure (combined result across N assays)

Power for difference test

© precisionbioassay inc

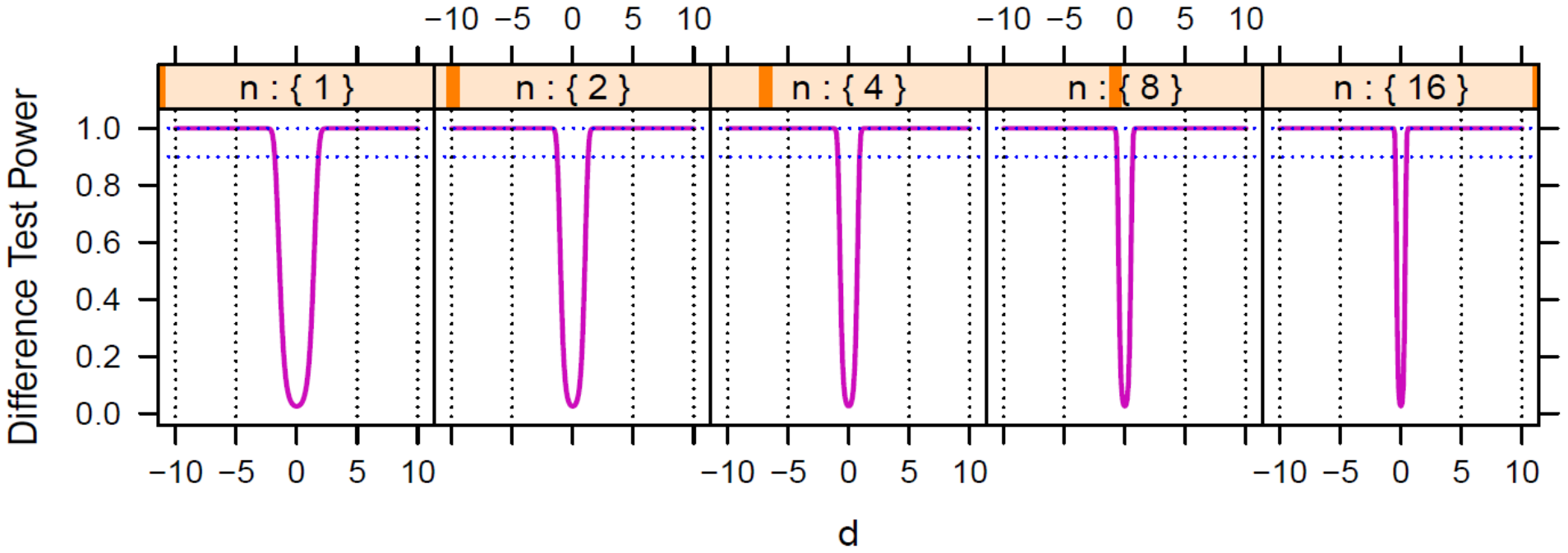
D. Lansky

$$\text{Power}_{\text{diff}} = T\left(\frac{nd^2}{s^2} - t_{1-\alpha_{\text{diff}},df}/2, df\right)$$

- where:
  - d is the true amount of non-similarity (one parameter's measure)
  - $\text{Power}_{\text{diff}}$ is P(correctly detecting true $d \neq 0$)
  - α_{diff} is the P(incorrectly declaring $d \neq 0$)
  - n is sample size (number of replicates of sample*dose within assay)
  - σ (estimated by s) is the SD of $\hat{d}$ (estimate of d)
  - T and t are t- CDF and PDF, indexed by df (in s)
- Power for difference test given (d) sensitive to: σ (s), α_{diff}, & n
- Difference test: under H_0: $d = 0$
- Note: because d always appears with s (or σ) can use 'new' $d = \frac{d}{\sigma}$

Difference test power

D. Lansky



- columns of panels for n
- horizontal dotted lines at power of .9 and 1
- vertical dotted lines at d of $\pm 5\sigma$ & 10σ

Power for equivalence test

D. Lansky

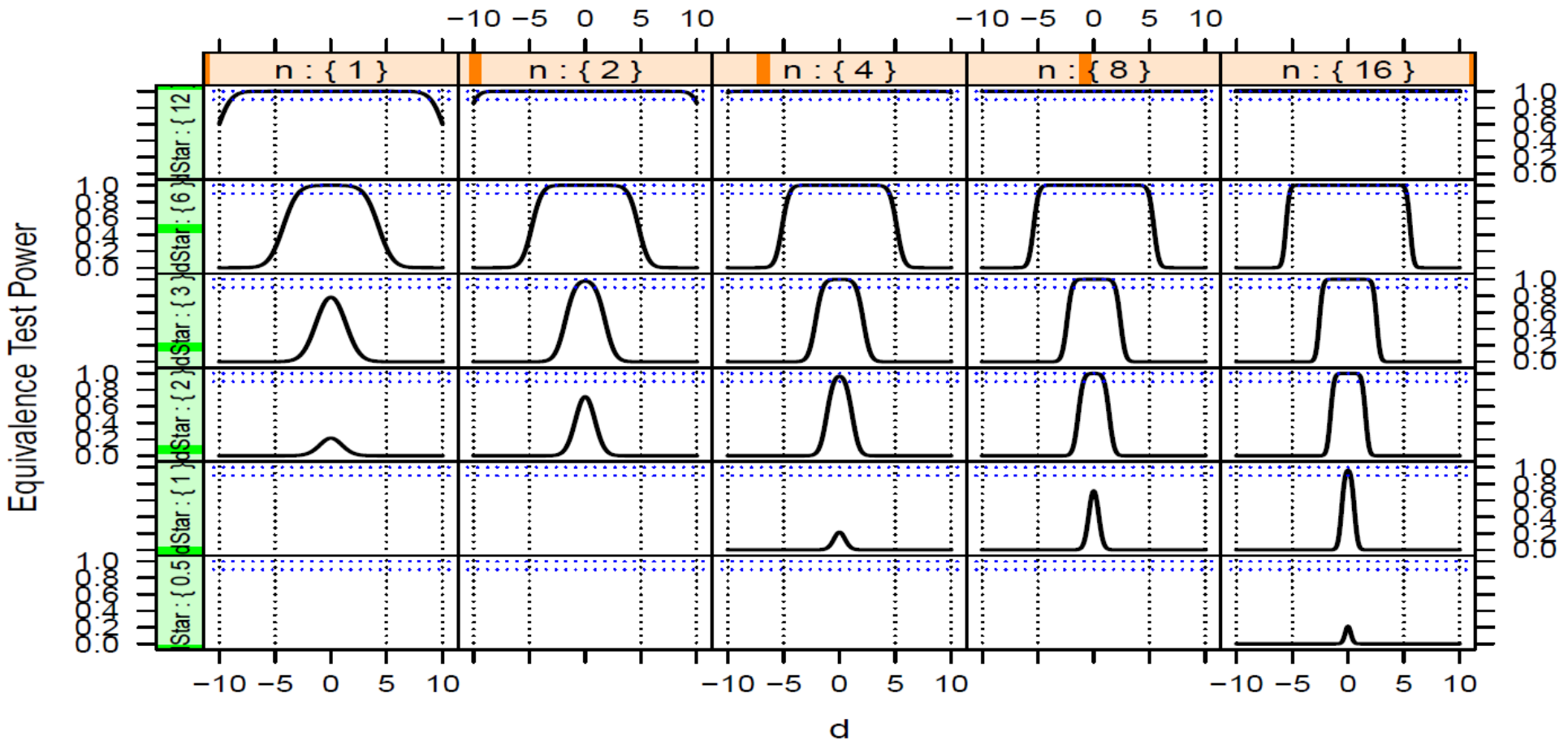
$$\text{Power}_{\text{equiv}} = T\left(\frac{d^* - d}{s/\sqrt{n}} - t_{1-\alpha_{\text{equiv}},df}, df\right) - T\left(\frac{-d^* - d}{s/\sqrt{n}} + t_{1-\alpha_{\text{equiv}},df}, df\right)$$

- where:
  - $\text{Power}_{\text{equiv}}$ is P(correctly detecting true $-d^* < d < d^*$)
  - α_{equiv} is the P(incorrectly declaring $-d^* < d < d^*$)
  - $\pm d^*$ are the equivalence test bounds
- Power for equivalence test (δ^*) sensitive to: σ, α_{diff}, & n, & d^*
- Equivalence test d under H_0: $d < -d^*$ or $d^* < d$

Equivalence test power

© precisionbioassay inc

D. Lansky



- rows of panels for $d^*\sigma$

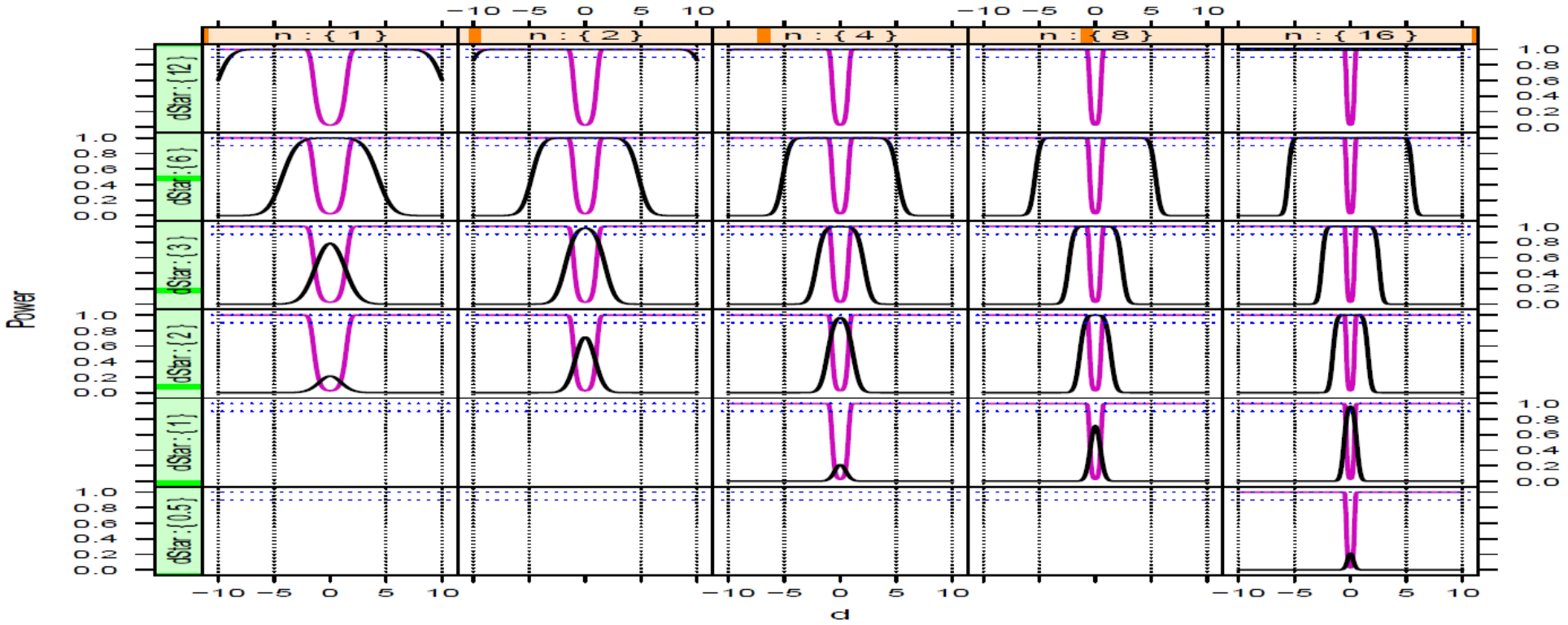
Bioassay purposes

D. Lansky

- Mature lot release => detecting non-similarity (diff test) less important
- Stability, qualify new process, etc. => difference test important
  - Limits on allowed non-similarity from potency bias limits
  - Assay capability (σ) and power needs determine n
- For all uses, limiting potency bias (or subject matter knowledge) limit may lead to narrow similarity limits (difference test or equivalence test?)
- Note: nonsimilarity of no-dose asymptote (D) isn't about sample (ref & test diluted to no-effect levels), it indicates an assay problem (cells, location, etc.)

Difference and equivalence test powers

D. Lansky



- is $d^* \approx 2\delta^*$ sensible?
- if equivalence power (for assay) too low, consider:
  - use difference test to protect potency against bias
  - perform equivalence test of similarity in procedure

Key messages

© precision bioassay inc

D. Lansky

- Need:
  - narrow limits on non-similarity to prevent bias in potency (diff test)
  - maybe narrower limits on some based on subject matter
  - wider limits (2x?) on non-similarity to pass similarity with equivalence tests
- Combining non-similarity across assays is sensible
  - Caution: don't simply repeat similarity fails (whether 'assay' or 'sample'); track
  - Difference tests of similarity at assay level (with assay size to have adequate power) wise to prevent (a largely unrecognized source of) potency bias
  - If method has poor similarity equivalence capability, do at procedure
- Interesting (hard) question: how to use (when combining results) results from samples (in OK assays) that fail difference test similarity?

Final thoughts

© precisionbioassay inc

D. Lansky

- Common issue: many bioassays - not enough info in an assay
  - To assess similarity well
  - To estimate uncertainty in similarity well
- Four ways to improve these estimates:
  - use mixed model analyses
  - pool similarity σ across samples
  - use historical info (about σ)
  - shift decision about similarity from method to procedure
- All 4 a challenge for in-house regulatory
- Consider using non-similarity diff test/assay and outlier (?) test on log potencies as part procedure for combining potencies
- What to monitor & how to adapt?
  - Monitor precision of (non)-similarity (σ) in assay
  - Adjust n in assay for power of non-similarity difference test
  - Adjust N assays in procedure for
    - non-similarity equivalence power &
    - precision of potency

Acknowledgements

- USP and USP bioassay panel members
- Carrie Wager
- Ramiro Barrantes-Reynolds
- Freyja Lynn
- Laureen Little
- NIH SBIR Phase I 'Better bioassays via designs for robots analyses with improved model selection and similarity bounds that limit potency bias' 1R43GM140743-01 (2021-2023)
- NSF EPSCoR 'Equivalence-based Random Effects Model Selection for Bioassay' OIA-1556770 subaward 03735U852412 (2017-2018)
- Many consulting clients