Recent results on equivalence test bounds to limit non-similarity induced bias in potency, with additional discussion about power

David Lansky, Ph.D.

precision bioassay inc

Burlington, Vermont, USA

Abstract

A major goal of bioassay development, validation, and monitoring is to minimize bias of potency. Testing for similarity via equivalence tests has become an essential part of modern bioassay analyses. Sensitivity analyses, reported here, show that scaled shifts in parameters measure non-similarity in ways that are assay-independent. We show that well-chosen similarity equivalence bounds limit bias in potency due to non-similarity. Hence, equivalence bounds for non-similarity can be informed by bias limits based on product specifications and the analytic target profile.

Why Test For Similarity?

- Fundamental assumption in bioassay
- Similarity of range and 'shape' support this assumption
- No-dose asymptote similarity very important
- Similarity of ALL non-EC50 parameters important
- Previous work showed: non-similarity causes potency bias

This talk will focus on:

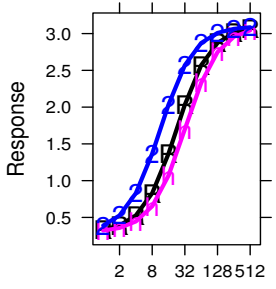
- Some types of non-similarity cause large bias in potency
- Each intended use of an assay has limits on potency bias
- Bias limits yield minimal similarity equivalence bounds
- Equivalence bounds $\pm d^*$ are not $\pm\delta^*$ (Berger & Hsu, 1996)
- Important to consider power for equivalence region
- A proposal to assess similarity on combined results

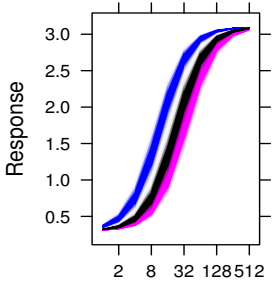
Bias/Precision Tradeoff?

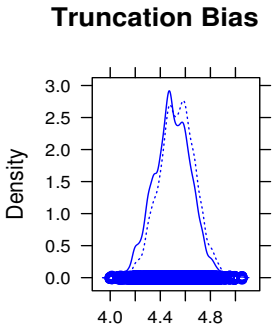
- Precision desirable, improves with n ($s \sim 1/\sqrt{n}$).
- Bias is a problem. Replication doesn't reduce bias.
- Bias limits for validation of an analytic method should be sensible given the product specs.
- Budget for (compared to room in specification):
  - manufacturing variance
  - measurement variance and n
  - degradation
  - all bias sources combined

Multiple Causes of Bias

- Measurable (and can be monitored):
  - Location/sequence bias
  - Truncation bias
  - Bias trend (across potency range)
- Hard to measure bias and only recently understood
  - Bias due to allowed non-similarity







The Problem

- Equivalence (now) non-controversial for similarity
- USP <1032> offers four ways to set equivalence bounds
- 3 with bounds based on assay capability
- Sensitivity (and robustness) of potency bias to non-similarity:
  - Visually
  - Quantitatively: % Geometric Bias of Potency $=$ $100\left(\text{anti-log}\left(\log\left(\hat{R}\right)-\log\left(R\right)\right)-1\right)$
- Similarity measures:
  - Parameter specific: scaled shifts
  - Composite: F, Chi^2, and scaled shift in A+D (Range+No-dose asymptote)

Intended Use Drives Requirements

- Narrow therapeutic window $\Rightarrow$ narrow potency specs (low bias & high precision) $\Rightarrow$ require an excellent assay
- Products w/narrow specs need narrow similarity bounds
- For all products, when qualifying new standard, production process, facility, or supplier of a critical reagent, low bias in potency may be required.

Four Parameter Logistic

$$y^* = \frac{A_i}{1 + e^{-B_i(\log(x) - C_i)}} + D_i + \epsilon$$

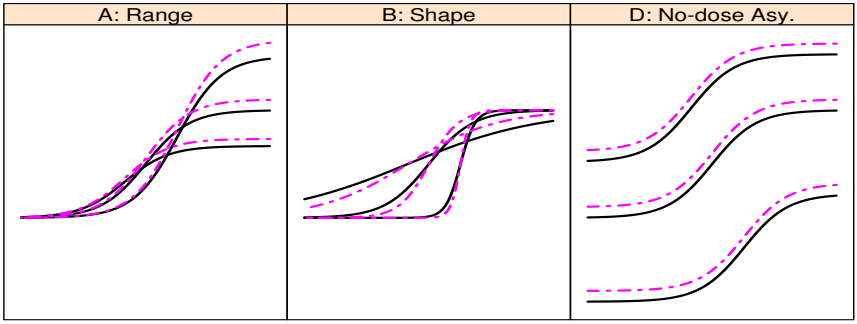
A = Response Range, B = "Shape",

C = Log EC50, and D = No-dose Asymptote (Ratkowsky & Reedy,1986)

Near Universal Scaled Shift Similarity Measures:

- $\%\Delta_A = 100 \times (A_{\text{Test}} - A_{\text{Ref}}) / \overline{A^*_{\text{Ref}}}$
- $\%\Delta_D = 100 \times (D_{\text{Test}} - D_{\text{Ref}}) / \overline{A^*_{\text{Ref}}}$ (Not a typo)
- $\%\Delta_B = 100 \times (B_{\text{Test}} - B_{\text{Ref}}) / \overline{B^*_{\text{Ref}}}$ (* Long term averages)

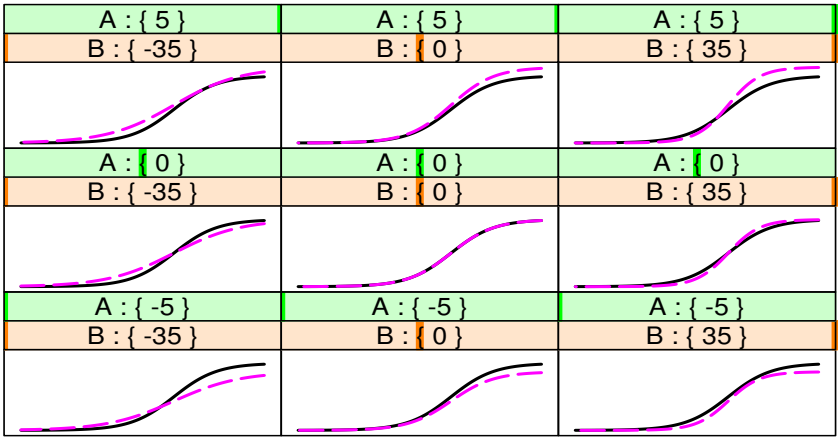
Scaled Shifts have consistent meaning



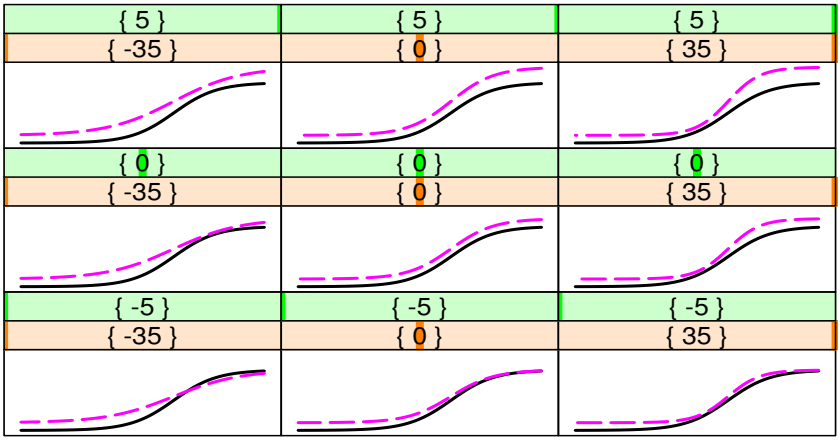
Black/Magenta pairs are standard/test: A and $D \times (2/3, 1, 3/2) + 10\%$, $B \times (1/3, 1, 3) + 50\%$

No-dose Asymptote 0% shift

D. Lansky



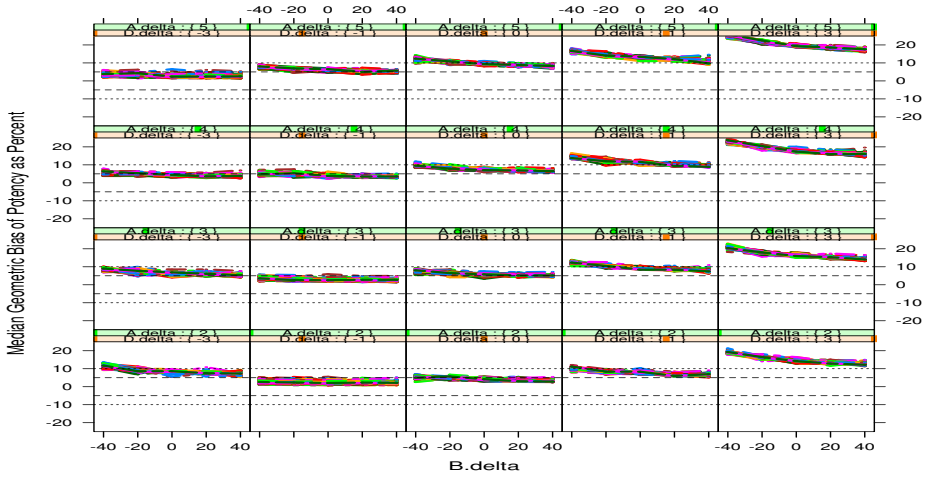
No-dose Asymptote +5% shift



Simulation Conditions

- potencies: 0.5, 1, 2
- resid SD: 2%, 3%, 4% of response range (A)
- A.delta: 2, 3, 4, 5% (of A)
- B.delta: -40, -20, 0, 20, 40% (of B)
- D.delta: -3, -1, 0, 1, 3% (of A)
- nDoses: 10, 12, 18
- E(N On Asymptote): 2, 3
- E(N On Asymptote) beyond 95% or 99% of A
- Fixed: A: 2, B: 1.5, C: 3, D: 2, n simulated assays: 50

Potency Bias from non-similarity at 1

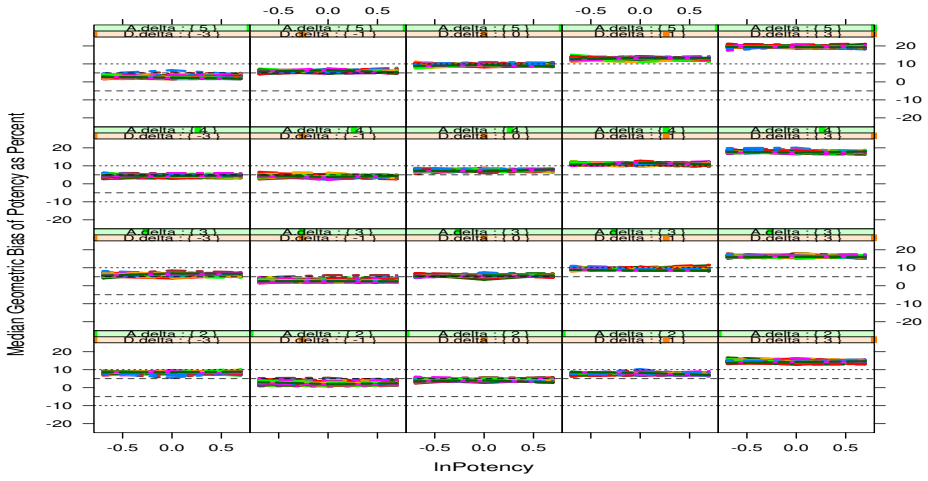


Robust to residual SD, nDoses, ENOnAsy, onAsy

Potency Bias w/B Similar due to other non-similarity

© precisionbioassay Plc

D. Lansky

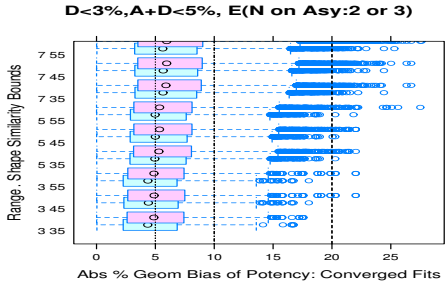


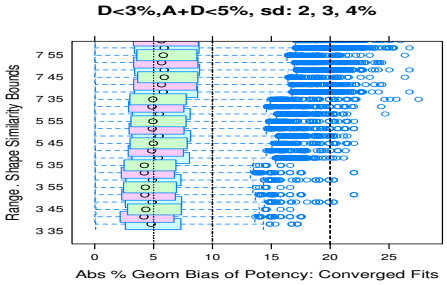
Robust to potency, residual SD, nDoses, ENOnAsy, onAsy

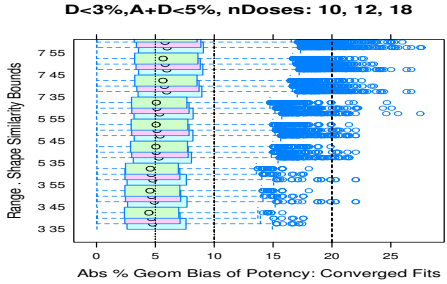
Potency Bias: dose range, n doses, sd, potency

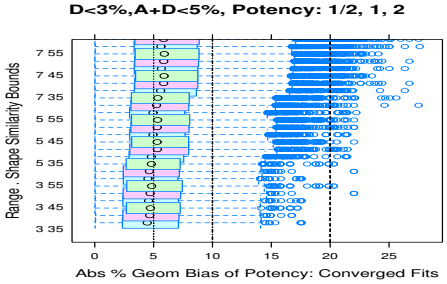
© precisionbioassay Plc

D. Lansky



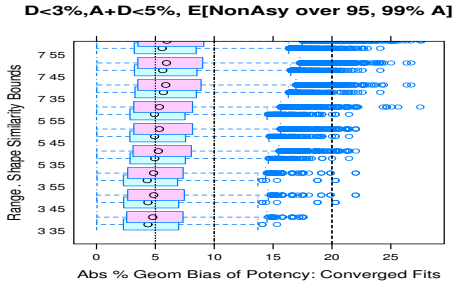


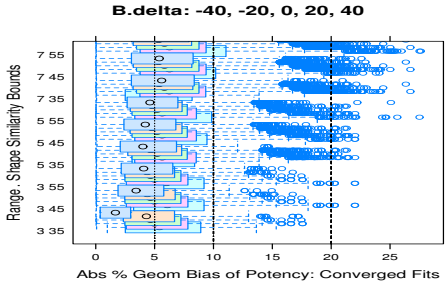


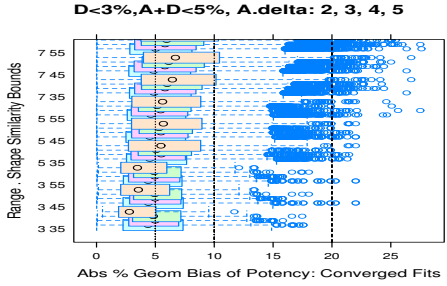


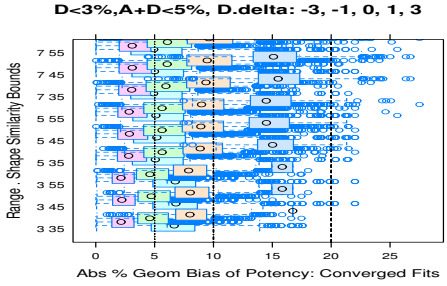
Potency Bias: dose range & nonSimilarity

D. Lansky



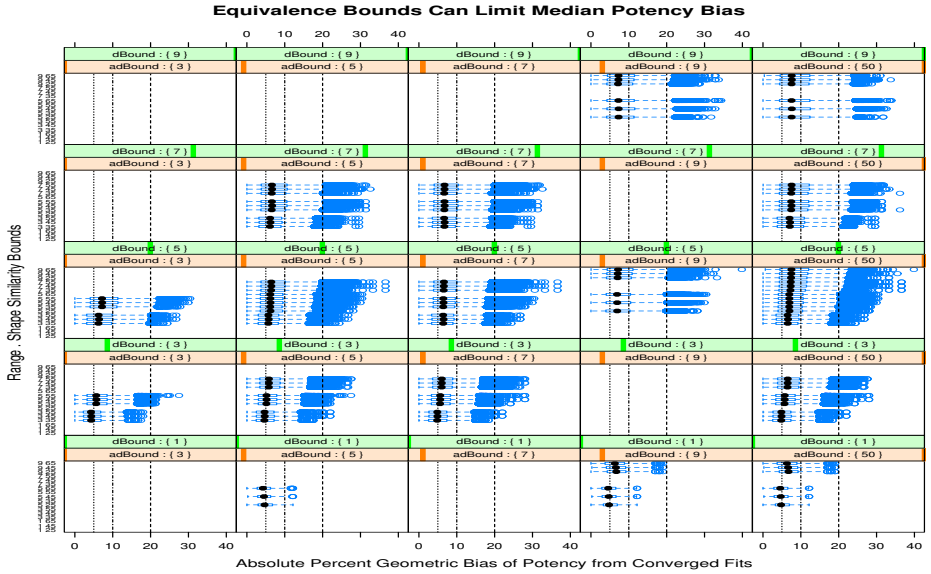






Potency Bias: Equivalence Bounds

D. Lansky



Results Summary

- Stringent convergence criteria important
- non-similarity (especially D) matters
- potency bias is skewed, with big range
- median bias of potency (where we should focus) can be controlled with equivalence bounds
- This bias is robust to:
  - assay design (n, E[n on asy])
  - some assay properties (A, B)
  - sample potency (0.5 to 2.0)
- BUT: narrow equivalence bounds appear needed

Power for an (as practiced) Equivalence Test

$$\text{Power} = P\left(\text{reject } H_0 | \delta\right)$$

$$= P\left(-d^* < \overline{d} - \frac{st_{1-\alpha,df}}{\sqrt{n}} \text{ and } \overline{d} + \frac{st_{1-\alpha,df}}{\sqrt{n}} < d^*\right)$$

$$= P\left(-d^* + \frac{st_{1-\alpha,df}}{\sqrt{n}} < \overline{d} < d^* - \frac{st_{1-\alpha,df}}{\sqrt{n}}\right)$$

with $\pm d^*$ the equivalence bounds, $d \sim N\left(\delta, \sigma^2\right)$, and $s = \text{SD}(d)$ and estimates σ.

Power (cont.)

Because $\overline{d} \sim N\left(\delta, \frac{\sigma^2}{n}\right)$ and $\frac{\sqrt{n}\left(\overline{d}-\delta\right)}{s} \sim t_{n-1}$:

$$= P\left(\frac{\sqrt{n}\left(-d^* - \delta\right)}{s} + t_{1-\alpha,df} < t_{df} < \frac{\sqrt{n}\left(d^* - \delta\right)}{s} - t_{1-\alpha,df}\right)$$

with $df = n - 1$.

Sample Size or δ for Equivalence Tests

D. Lansky

USP <1033> says:

$$n \geq \frac{\left(t_{\alpha,df} + t_{\beta/2,df}\right)^2 \sigma^2}{(\delta)^2}$$

Which should give:

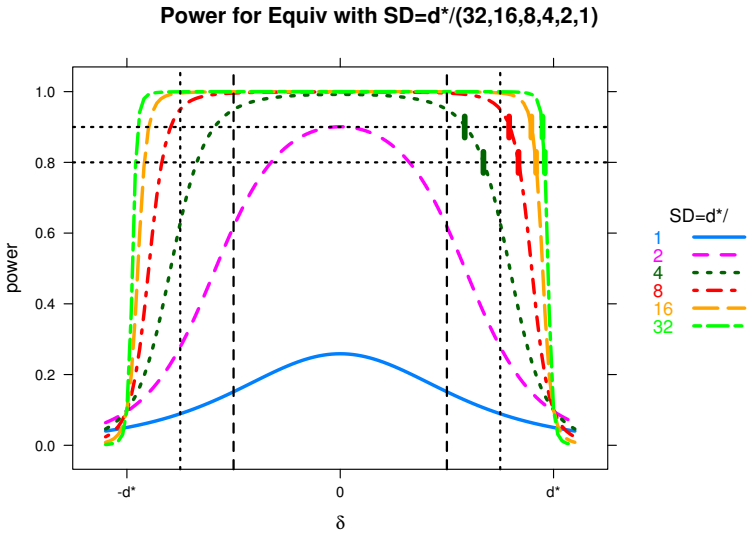
$$\delta \geq \sqrt{\frac{\left(t_{\alpha,df} + t_{\beta/2,df}\right)^2 \sigma^2}{n}}$$

I get something that involves d^*:

$$d^* - \delta \geq \frac{\sigma}{\sqrt{n}} \left(t_{1-\frac{\beta}{2},df} + t_{1-\alpha,df}\right)$$

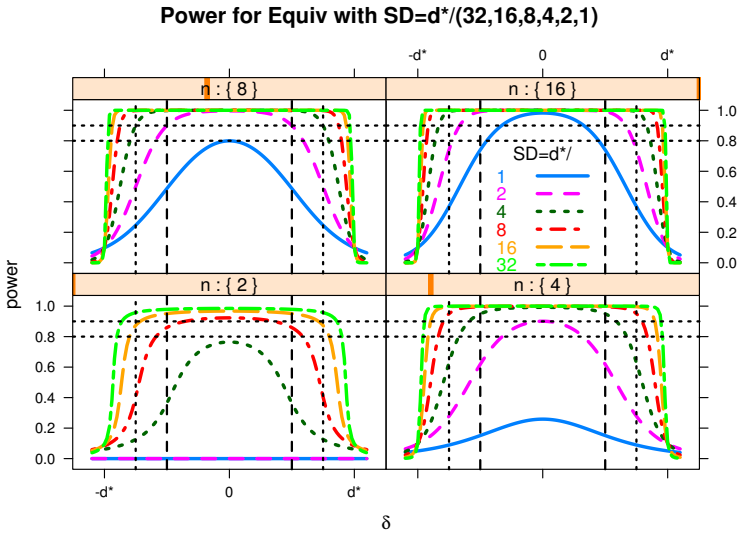
Power for Equivalence Tests with n=4, d^*

D. Lansky



Power for Equivalence Tests Using d^*

D. Lansky



Discussion

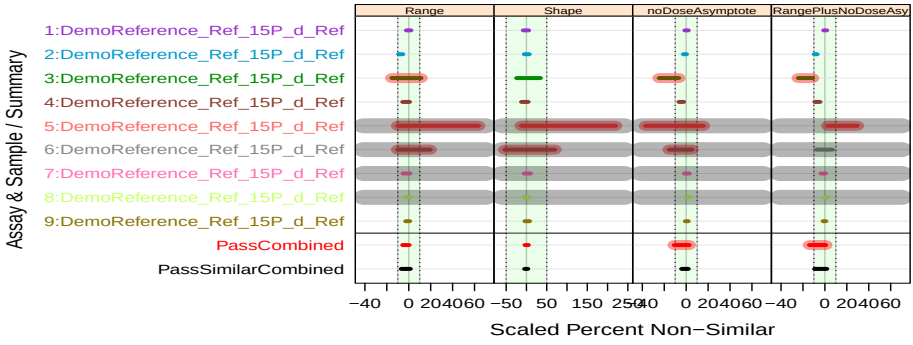
- Equiv bounds to limit potency bias: a minimal requirement
- Other targets for sensitivity?
- Allowable non-similarity potency bias: bias budget
- Experience: modest equivalence bounds are demanding
- USP methods a & b for setting equiv bounds use $\delta = 0$
- Sensible to require CI for non-similarity shorter than d^*
- Power to reliably pass sensible bias-protective equivalence bounds will require very good assays or ...

Proposal

- Relax per-assay equiv. bounds OR outlier test potencies
- Impose stringent equivalence bounds on combined results across replicate assays (to prevent bias and make inference to lot)
- Require good power for 50%? of similarity region (narrow equivalence CI)
- This approach will require more assay replicates

Combining Equivalence Across Assays

D. Lansky



Acknowledgements

- Consulting clients
- USP and USP bioassay panel members
- Carrie Wager
- Ramiro Barrantes
- Mark Blanchard
- NSF EPSCoR 30373SUB52412
- NIH SBIR 3R44RR02198-03S1

Your turn

- Questions?