

Parallelism Testing Methods for Potency Bioassay Curves*

Meiyu Shen, PhD

Office of Biostatistics, Center for Drug Evaluation and
Research, FDA

Disclaimer: This presentation reflects the views of the authors and should not be construed to represent FDA's views or policies.

Ji Li , Chang Chen , Meiyu Shen & Yi Tsong (2026) Parallelism of four-parameter logistic bioassay models: equivalence testing for parameter differences, Journal of Biopharmaceutical Statistics, 36:4, 583-599, DOI: 10.1080/10543406.2026.2670515

What is Biological Assay?

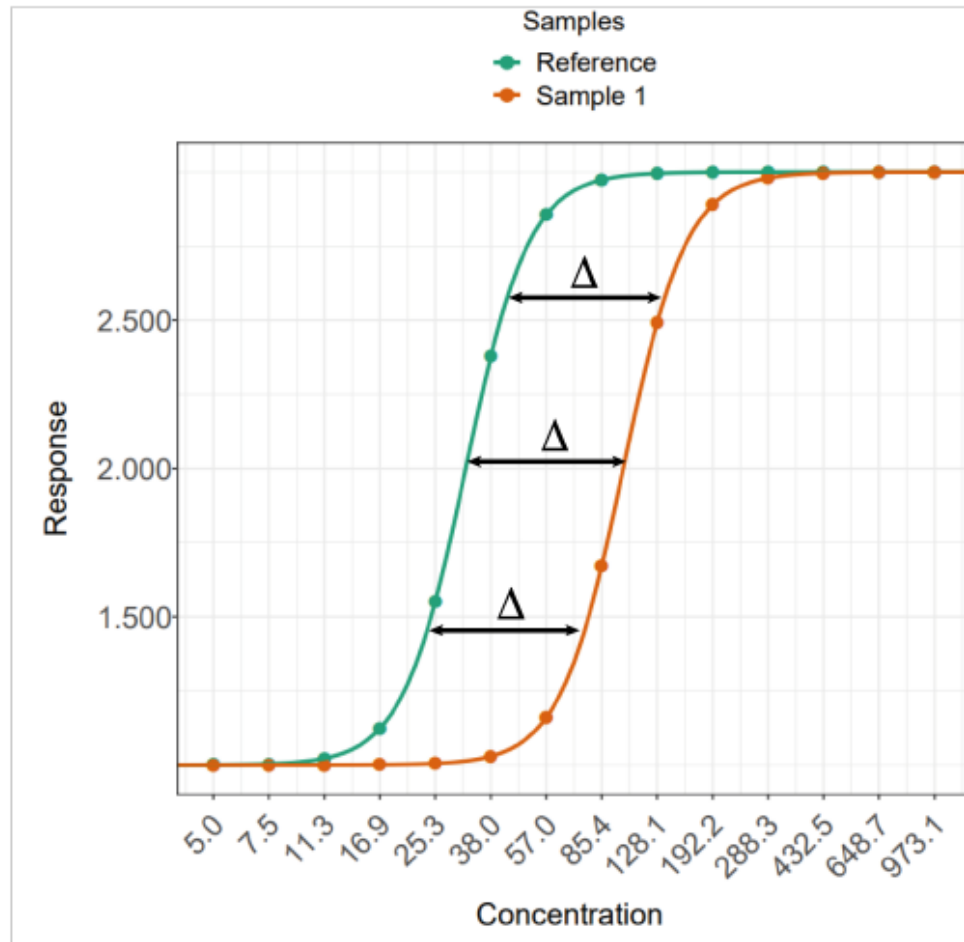
- Biological assays, or bioassays
 - a measurement method capable of quantifying a response from living matter (e.g., cells, tissues, and animals) that is treated with some stimulus (e.g., therapeutic preparation, prophylactic preparation, or toxin).
- Relative potency
 - A biological activity or potency, often obtained by comparing the sample under study with a standard.
 - Given biological similarity
 - the standard and the test sample should contain comparable active material to make the comparison meaningful

Two conditions for determining relative potency

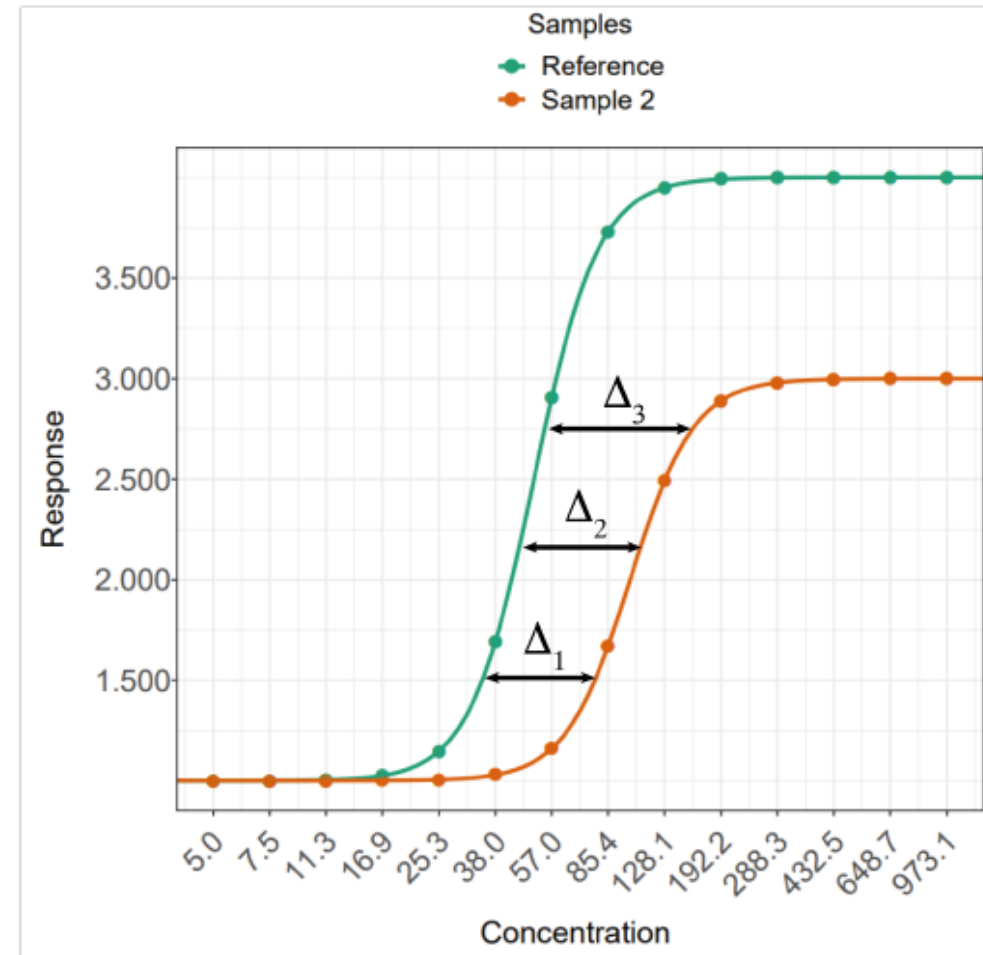
- To determine relative potency of the test versus the standard from an indirect bioassay
 - Condition 1
 - Confirm biological similarity between Test and Standard
 - Condition 2
 - Address how similar the dose-response curve of the Test is compared to that of the Standard by Parallelism testing methods given that a biological similarity exists between the Test and the Standard

Model for dose-response curve:
$$f(x|\theta_i) = a_i + \frac{b_i - a_i}{1 + \exp(d_i(x - \log(c_i)))}$$

where $\theta_i = (a_i, b_i, c_i, d_i)'$ is the parameter vector.



(a) Two parallel curves, the horizontal shift is the same for all doses



(b) Non-parallel curves

Evaluating parallelism in bioassay curves



- Reasonable similarity on :
 - no-dose asymptote, a
 - Maximum asymptote, b
 - Slope d
- The p-value method for difference testing $H_0: a_T = a_R; b_T = b_R; d_T = d_R$
 - Here T is for test and R is for Reference standard
 - Large p-value means that two curves are parallel.
 - The significance level of 0.05 is a common choice.
- Equivalence Testing for parallelism (more on next page).
 - Challenge: equivalence margin determination

Evaluating parallelism of bioassay curves



- Equivalence test for parallelism

- For a (no dose asymptote):

- $H_0: \frac{a_T}{a_R} \leq \delta_{aL} \cup \frac{a_T}{a_R} \geq \delta_{aU}; H_a: \delta_{aL} < \frac{a_T}{a_R} < \delta_{aU}$

(1)

- For b (maximum asymptote):

- $H_0: \frac{b_T}{b_R} \leq \delta_{bL} \cup \frac{b_T}{b_R} \geq \delta_{bU}; H_a: \delta_{bL} < \frac{b_T}{b_R} < \delta_{bU}$

- For d (slope):

- $H_0: \frac{d_T}{d_R} \leq \delta_{dL} \cup \frac{d_T}{d_R} \geq \delta_{dU}; H_a: \delta_{dL} < \frac{d_T}{d_R} < \delta_{dU}$

- Equivalence margins.

- Jonkman and Sidik (2009) proposed (0.8, 1.25) as the margin for all three ratios

- Hauck et al. (2005) proposed some data-driven equivalence margins

Four scenarios of true test and reference bioassay curves

Table 2. Monte Carlo simulation scenarios.

Scenario	Reference standard				Sample			
	a_1	b_1	c_1	d_1	a_2	b_2	c_2	d_2
1	2	10	30	1.25	2	10	30	1.25
2	2	10	30	1.25	2.02	10.26	60	1.33
3	2	10	30	1.25	1.6	8	30	1.56
4	2	10	30	1.25	2	15	60	2

- Scenario 1 represents identical curves;
- Scenario 2 represents two parallel equivalent curves by definitions in Eq. (1);
- Scenario 3 represents two borderline parallel curves by definition in Eq. (1);
- Scenario 4 represents non-parallel curves.

Equivalence test for difference in parameters



- Equivalence test for parallelism
 - For a (no dose asymptote):
 - $H_0: a_T - a_R \leq \delta_{aL} \cup a_T - a_R \geq \delta_{aU}; H_a: \delta_{aL} < a_T - a_R < \delta_{aU}$
 - For b (maximum asymptote):
 - $H_0: b_T - b_R \leq \delta_{bL} \cup b_T - b_R \geq \delta_{bU}; H_a: \delta_{bL} < b_T - b_R < \delta_{bU}$
 - For d (slope):
 - $H_0: d_T - d_R \leq \delta_{dL} \cup d_T - d_R \geq \delta_{dU}; H_a: \delta_{dL} < d_T - d_R < \delta_{dU}$
- Margin can be derived using tolerance interval based on historical data under normality assumption

Simulate Historical data for deriving parallelism margins for equivalence test for difference in parameters

- Simulate historical data
 - Reference standard: $a=2$, $b=10$, $c=30$, $d=1.25$.
 - Test: 26 cases in each scenario excluding ($a = 2$, $b = 10$, $c = 30$, $d = 1.25$)

Test Scenario	a	b	c	d
1, non-parallel	$2*0.7$, 2, $2/0.7$	$10*0.7$, 10, $10/0.7$	30	$1.25*0.8$, 1.25, $1.25/0.8$
2, parallel	$2*0.8$, 2, $2/0.8$	$10*0.8$, 10, $10/0.8$	30	$1.25*0.8$, 1.25, $1.25/0.8$
3, parallel	$2*0.9$, 2, $2/0.9$	$10*0.9$, 10, $10/0.9$	30	$1.25*0.9$, 1.25, $1.25/0.9$
8, parallel	$2*0.95$, $2/0.95$	$10*0.95$, $10/0.95$	30	$1.25*0.95$, $1.25/0.95$
4, parallel	$2*0.99$, $2/0.99$	$10*0.99$, $10/0.99$	30	$1.25*0.99$, $1.25/0.99$

Parametric methods for equivalence margins



- We propose
 - Following steps are for parameter a.
 - Obtain $\hat{a}_{1T}, \hat{a}_{2T}, \dots, \hat{a}_{NT}$ and $\hat{a}_{1R}$ from fitting $f(\theta_i, x) = a_i + \frac{(b_i - a_i)}{1 + \exp\{d_i(x - \log(c_i))\}}$ to
 - Each of N dose-response curves of historical data for test.
 - 1 dose-response curve of historical data for reference standard.
 - Calculate CIs of differences of parameters between each of N historical datasets and one reference standard data.
 - Upper tolerance limit (TL) = mean(upper bound) + k*s; lower TL = mean(lower bound) - k*s, where k is a tolerance factor, s is the SD.
 - Equivalence limit = max(abs(lower TL), abs(upper TL)).
 - Repeat above steps for b and d.

SIMULATION SETUP

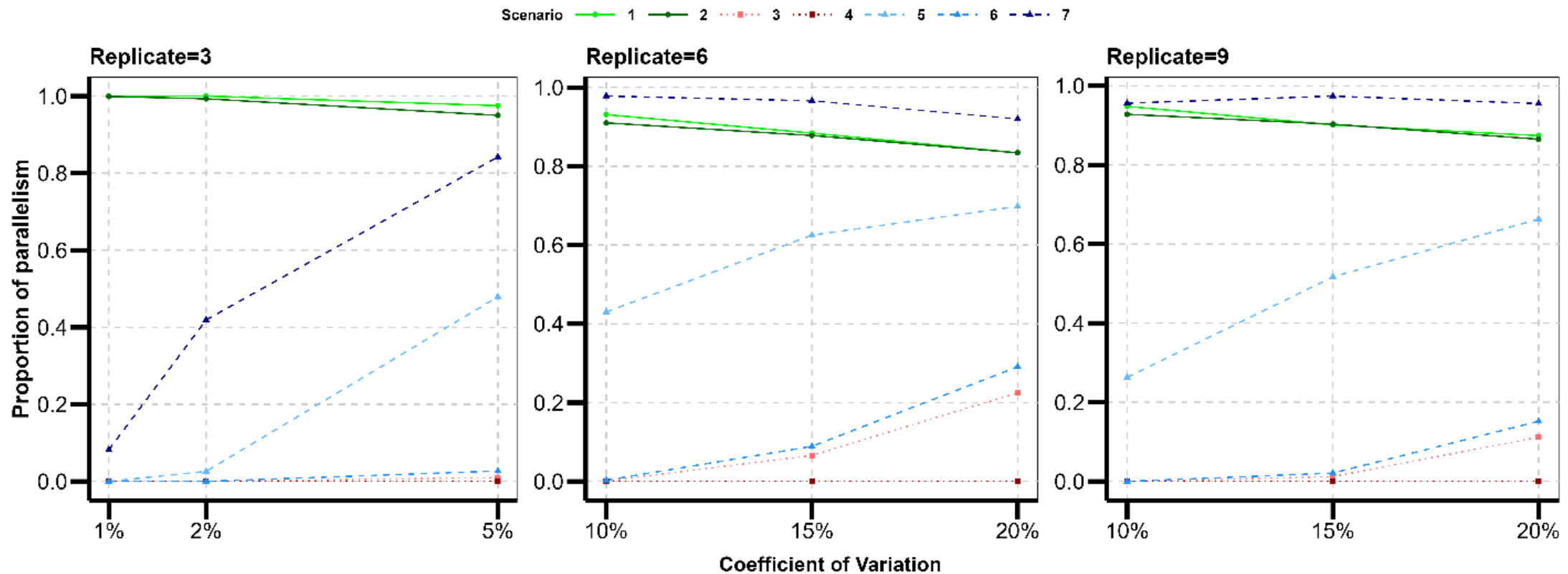
	Scenario	Reference standard				Sample			
		a_1	b_1	c_1	d_1	a_2	b_2	c_2	d_2
Parallel	1	2	10	30	1.25	2	10	30	1.25
	2	2	10	30	1.25	2.02	10.26	60	1.33
Nonparallel	3	2	10	30	1.25	1.6	8	30	1.56
	4	2	10	30	1.25	2	15	60	2
Borderline parallel	5	2	10	30	1.25	1.6	10	30	1.25
	6	2	10	30	1.25	2	8	30	1.25
	7	2	10	30	1.25	2	10	30	1

Consider 8 concentrations for each curve. For each concentration, replicates=3 when CV=1% and 5%; replicates=6 when CV=10% and 15%; replicates=9 when CV=20%

Simulation results: margins derived from historical dataset 3 (parallel, test/ref=0.9) with parametric method



Parametric Method for Historical Scenario 3 Across Seven Test Scenarios



1, 2: parallel, 3,4: nonparallel, 5,6,7: borderline parallel

Conclusions

- Reasonable equivalence margins can be determined from historical data if the reference and test in historical data are approximately parallel, e.g., historical data 3 (90% of the standard), and 8 (95% of the standard)
- The equivalence limits should not be determined from historical data if the reference and test in historical data are not parallel
- Too stringent equivalence margins are derived from historical data if the reference and test in historical data are almost perfectly parallel

References

- Jeffrey N Jonkman and Kurex Sidik. “Equivalence testing for parallelism in the four-parameter logistic model”. In: Journal of biopharmaceutical statistics 19.5 (2009), pp. 818–837.
- Walter W Hauck et al. “Assessing parallelism prior to determining relative potency”. In: PDA Journal of Pharmaceutical Science and Technology 59.2 (2005), pp. 127–137.

Thank you

