Comparison of Methods Used to Measure Parallelism

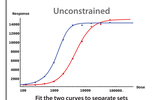
John R. Dunn II, Ph.D. Chief Technical Officer, Brendan Technologies Inc., Carlsbad, CA, email: jdunn@brendan.com

BRENDAN TECHNOLOGIES
Powerful Analysis. True Results.

Introduction

There are three methods currently in use to measure the similarity of the dose response curves of two samples. Two of these methods utilize the extra sum of squares test to compute the difference in the residuals between the constrained regression (same shape for both samples) and their unconstrained regressions. One method uses an accurately weighted full curve regression and chi-square test, and the other uses a parallel line regression and F test. The third method uses an equivalence approach that compares the confidence limits of the regression parameters. Data from samples with identical material to the reference standards and samples that have been stressed or have impurities are examined using all three methods.

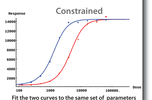
Parallelism of Residuals (Extra Sum of Squares)



Fit the two curves to separate sets of parameters, and add their RSSEs.

$$SSE_{full}(\mathbf{p}_{ref}, \mathbf{p}_{alt}) = \sum_{i=1}^{N} w_i^{ref}\left(y_i^{ref} - f\left(x_i^{ref}, \mathbf{p}_{ref}\right)\right)^2 + \sum_{i=1}^{N} w_i^{alt}\left(y_i^{alt} - f\left(x_i^{alt}, \mathbf{p}_{alt}\right)\right)^2$$

$$RSSE_{unconst} = \min_{\mathbf{p}_{ref}, \mathbf{p}_{alt}} SSE_{unconst}(\mathbf{p}_{ref}, \mathbf{p}_{alt})$$



Fit the two curves to the same set of parameters (with an additional parameter r for the unknown relative potency) and add the RSSEs.

$$SSE_{red}(\mathbf{p}) = \sum_{i=1}^{N} w_i^{ref}\left(y_i^{ref} - f\left(x_i^{ref}, \mathbf{p}\right)\right)^2 + \sum_{i=1}^{N} w_i^{alt}\left(y_i^{alt} - f\left(x_i^{alt}, r \cdot \mathbf{p}\right)\right)^2$$

$$RSSE_{const} = \min_{r, \mathbf{p}} SSE_{const}(r, \mathbf{p})$$

Chi-Square (χ^2) Test

$$RSSE_{nonparallel} = RSSE_{constrained} - RSSE_{unconstrained}$$

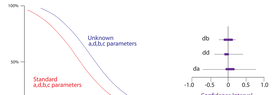
- The $RSSE_{nonparallel}$ is an actual measure of parallelism. It is not just a p-value "Yes" / "No" result.
- The $RSSE_{nonparallel}$ is zero if the two curves are parallel.
- The $RSSE_{nonparallel}$ increases with nonparallelism.
- When the data is accurately weighted, the $RSSE_{nonparallel}$ is distributed as a chi-square variable with P-1 degrees of freedom.
- The $RSSE_{nonparallel}$ (X^2 statistic) and X^2 probability allow a threshold of acceptable nonparallelism to be determined empirically.

Parallel Line Assay (F Test)

$$\text{F Statistic} = \frac{\dfrac{RSSE_{nonpar}}{df_{nonpar}}}{\dfrac{RSSE_{unconst}}{df_{unconst}}} = \frac{\dfrac{RSSE_{const} - RSSE_{unconst}}{df_{const} - df_{unconst}}}{\dfrac{RSSE_{unconst}}{df_{unconst}}}$$

- The F Statistic and F Probability can be distorted by the fit of the unconstrained model, and do not measure directly the amount of nonparallelism.
- This method is often used because it does not absolutely require correct weighting, since the errors in the $RSSE_{nonparallel}$ due to incorrect weighting are compensated by the errors in the $RSSE_{unconstrained}$.

Equivalence Test



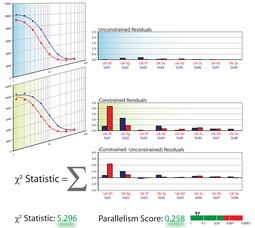
$$\text{Ln}(\text{Response}) = (D - S_i * dD) + \frac{(A - S_i * dA) \cdot (D - S_i * dD)}{1 + \text{Exp}(B - S_i * dB) * (LC - S_i * dC - \text{Ln}(\text{Concentration}))}$$

- (SL Concentration, Response) are combined data sets (Std and Uk)
- S_i = 1 for Std data points and 0 for Uk data points
- A, D, B and LC are the 4PL curve parameters for the Uk curve fit
- (dA, dD, dB and dLC are the respective parameter differences between the Uk and Std curve fits

150% Potency Control

5PL, Chi-Square

Relative Potency: 1.599

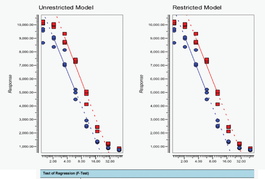


$$\chi^2 \text{ Statistic} = \sum$$

χ^2 Statistic: 5.296 Parallelism Score: 0.258

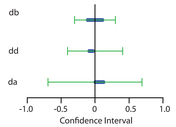
Parallel Line, F-Test

Relative Potency: 1.675



Test of Regression (F-Test)			
Parallel: 276.758		This test passes $F_{Parallel} < F_{Crit}$	Test Passed
Fdf(1,5): 7.158			
Test of Linearity (F-Test)			
Fdf(1,5): 3.541	5.916	This test passes $F_{Linearity} < F_{Crit}$	Test Passed
Test of Parallelism (F-Test)			
Fdf(1,5): 6.168	5.916	This test passes $F_{Parallelism} < F_{Crit}$	Test Passed

4PL, Equivalence

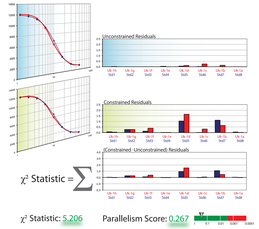


Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL
a	6.4908331792	0.0160628	6.4587398	6.5230263
d	9.1824336538	0.01123147	9.16196489	9.20360486
b	1.9818224621	0.03971267	1.90934852	2.05507032
c	2.0152204835	0.02247974	2.32745205	1.70298765
da passed	0.0632596881	0.00245571	0.06034328	0.0661839
dd passed	-0.057632479	0.01450578	-0.0847167	-0.0307323
db passed	-0.051877459	0.05851344	-0.1069601	0.11046596
dc	-0.44838624	0.05966522	-0.4796278	-0.4178319

100% Potency Control

5PL, Chi-Square

Relative Potency: 1.041

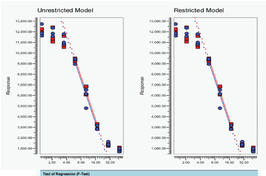


$$\chi^2 \text{ Statistic} = \sum$$

χ^2 Statistic: 5.206 Parallelism Score: 0.267

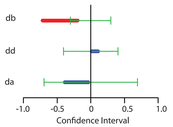
Parallel Line, F-Test

Relative Potency: 1.058



Test of Regression (F-Test)			
Parallel: 573.607		This test passes $F_{Parallel} < F_{Crit}$	Test Passed
Fdf(1,5): 7.158			
Test of Linearity (F-Test)			
Fdf(1,5): 1.074	5.916	This test passes $F_{Linearity} < F_{Crit}$	Test Passed
Test of Parallelism (F-Test)			
Fdf(1,5): 0.001	5.916	This test passes $F_{Parallelism} < F_{Crit}$	Test Passed

4PL, Equivalence



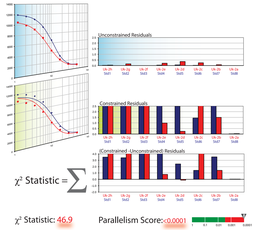
Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL
a	6.3815952743	0.0798509	6.2173766	6.51427418
d	9.2674048106	0.02619901	9.2203663	9.31886673
b	1.9093056416	0.0516827	1.66400789	2.06818246
c	2.1192058835	0.0334876	2.86869134	2.39393084
da passed	0.0323024303	0.02238997	-0.12798487	0.0029599
dd passed	-0.103720706	0.03301365	-0.1648954	-0.0404282
db failed	-0.501430687	0.13737399	-0.8184687	-0.1881856
dc	-0.104392766	0.0473757	-0.1892034	-0.0143556

- The equivalence method falsely flagged this reference material as being nonparallel.

Isoform Impurity

5PL, Chi-Square

Relative Potency: 1.356

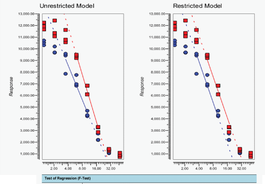


$$\chi^2 \text{ Statistic} = \sum$$

χ^2 Statistic: 46.9 Parallelism Score: <0.0001

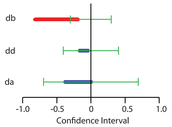
Parallel Line, F-Test

Relative Potency: 1.545



Test of Regression (F-Test)			
Parallel: 748.931		This test passes $F_{Parallel} < F_{Crit}$	Test Passed
Fdf(1,5): 7.158			
Test of Linearity (F-Test)			
Fdf(1,5): 1.400	5.916	This test passes $F_{Linearity} < F_{Crit}$	Test Passed
Test of Parallelism (F-Test)			
Fdf(1,5): 0.000	5.916	This test passes $F_{Parallelism} < F_{Crit}$	Test Passed

4PL, Equivalence



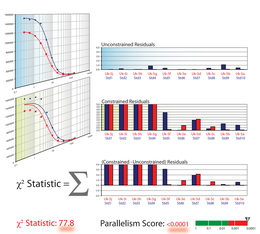
Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL
a	6.3829542743	0.0798509	6.2173766	6.51427418
d	9.2674048106	0.02619901	9.2203663	9.31886673
b	1.8631516827	0.1091548	1.66400789	2.06818246
c	2.1192058835	0.0334876	2.86869134	2.39393084
da passed	0.0323024303	0.02238997	-0.12798487	0.0029599
dd passed	-0.103720706	0.03301365	-0.1648954	-0.0404282
db failed	-0.501430687	0.13737399	-0.8184687	-0.1881856
dc	-0.104392766	0.0473757	-0.1892034	-0.0143556

- The Chi-Square method is capable of detecting nonparallelism in the region spanning the 3 lowest concentration standards.
- The 3 lowest concentration standards are not part of the analyzed region thus the impurities escape detection when using parallel lines .

UV Degraded

5PL, Chi-Square

Relative Potency: 1.612

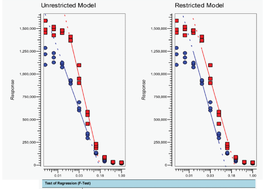


$$\chi^2 \text{ Statistic} = \sum$$

χ^2 Statistic: 77.8 Parallelism Score: <0.0001

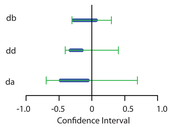
Parallel Line, F-Test

Relative Potency: 2.134



Test of Regression (F-Test)			
Parallel: 1290.068		This test passes $F_{Parallel} < F_{Crit}$	Test Passed
Fdf(1,5): 0.674			
Test of Linearity (F-Test)			
Fdf(1,5): 2.224	5.974	This test passes $F_{Linearity} < F_{Crit}$	Test Passed
Test of Parallelism (F-Test)			
Fdf(1,5): 56.058	5.974	This test passes $F_{Parallelism} < F_{Crit}$	Test Failed

4PL, Equivalence



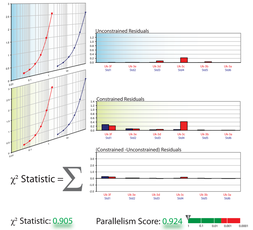
Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL
a	9.6354952377	0.0823461	9.4787182	9.7738627
d	14.01878902	0.09846213	13.9553329	14.0858123
b	1.9093056416	0.0516827	1.66400789	2.06818246
c	2.8763646654	0.0432587	2.80268529	2.96564605
da passed	-0.267495092	0.11931236	-0.477097	-0.0583189
dd passed	-0.238276408	0.04607462	-0.325001	-0.1505584
db passed	-0.106606034	0.09696685	-0.2808548	0.06616517
dc	-0.173909706	0.060071	-0.277809	-0.0657356

- The Chi-Square method is capable of detecting significant nonparallelism in the low concentration region.
- The Equivalence method failed to detect the UV degradation.

Accepted Vaccine Batch

5PL, Chi-Square

Relative Potency: 60.8

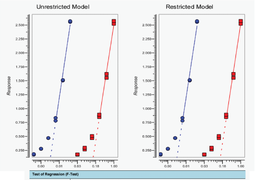


$$\chi^2 \text{ Statistic} = \sum$$

χ^2 Statistic: 0.905 Parallelism Score: 0.924

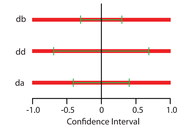
Parallel Line, F-Test

Relative Potency: 61.9



Test of Regression (F-Test)			
Parallel: 8636.248		This test passes $F_{Parallel} < F_{Crit}$	Test Passed
Fdf(1,5): 5.876			
Test of Linearity (F-Test)			
Fdf(1,5): 31.092	8.042	This test passes $F_{Linearity} < F_{Crit}$	Test Failed
Test of Parallelism (F-Test)			
Fdf(1,5): 2.985	13.745	This test passes $F_{Parallelism} < F_{Crit}$	Test Passed

4PL, Equivalence



Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL
a	2.9290032648	0.0627485	1.49669785	1.44813815
d	-3.037859936	0.4967773	.	-2.4335391
b	3.9924931364	0.61243355	.	0.84773714
c	3.0024491187	0.63132282	3.2857799	.
da failed	0.7464252761	1.60096681	-12.126619	.
dd failed	-0.17931544	0.70875465	.	.
db failed	-0.088648108	0.25658573	.	.
dc	0.9224414187	0.63132282	.	.

- The Parallel Lines method could not find a linear region.
- The Equivalence method falsely flagged this reference material.

Parallelism Analysis Results Comparison

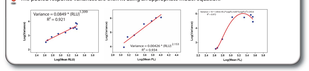
Assay	Sample	Material Tested	Parallelism Test Outcome			
			Expected Result	5PL, Chi-Square	Lin-Reg, F-Test	4PL, Equiv. Test
AED-12 Receptor-ligand binding inhibition assay	Uk-1	Reference standard	Pass	Pass	Pass	Fail
	Uk-2	Reference standard	Pass	Pass	Pass	Fail
	Uk-3	Reference standard	Pass	Pass	Pass	Pass
AED-15 Receptor-ligand binding inhibition assay	Uk-1	Reference standard	Pass	Pass	Pass	Fail
	Uk-2	Structural isoform impurity	Fail	Fail	Pass	Fail
	Uk-3	Structural isoform impurity	Fail	Fail	Pass	Fail
AED-75 Receptor-ligand binding inhibition assay	Uk-1	Reference standard	Pass	Pass	Pass	Fail
	Uk-2	Structural isoform impurity	Fail	Fail	Uncertain	Fail
	Uk-3	Structural isoform impurity	Fail	Fail	Fail	Uncertain
AM8-100 Cell proliferation assay	Uk-1	Reference standard	Pass	Pass	Pass	Pass
	Uk-2	UV degraded	Fail	Fail	Pass	Fail
	Uk-3	UV degraded	Fail	Fail	Pass	Fail
AM8-181 Cell proliferation assay	Uk-1	Reference standard	Pass	Pass	Pass	Pass
	Uk-3	UV degraded	Fail	Fail	Pass	Fail
RN3-10 Vaccine antigen assay	Uk-1	Reference standard	Pass	Pass	Uncertain	Fail
	Uk-3	Rejected vaccine batch	Fail	Fail	Pass	Fail
RN3-18 Vaccine antigen assay	Uk-1	Reference standard	Pass	Pass	Fail	Fail
	Uk-3	Rejected vaccine batch	Fail	Fail	Pass	Fail
RN3-19 Vaccine antigen assay	Uk-1	Reference standard	Pass	Pass	Uncertain	Fail
	Uk-2	Accepted vaccine batch	Pass	Pass	Uncertain	Fail
	Uk-3	Accepted vaccine batch	Pass	Pass	Uncertain	Fail

Why is Weighting Significant?

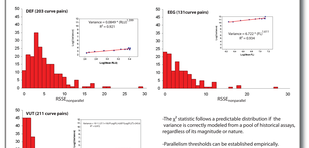
- Weighting the responses with the inverse of their variance is required for all regression fits to produce the Maximum Likelihood Estimate of the underlying curve.
- Allows each point to contribute equally to the fitted curve and prevents the curve being fitted predominantly to the high response data points. Bioassay and immunoassay data are very heteroscedastic, usually 2-4 orders of magnitude, and the variance profiles vary substantially between test methods.

Determining the Variance Model

- The underlying variance profile of the responses for a test method can be determined with pooled assays. The pooled assays will also incorporate the variation in precision observed between assays. The number of assays needed for a reliable estimate of the variance profile is dependent upon the precision of the responses.
  - -Dilution Replicates. The replicate variance of each dilution is obtained from pooled assays using the within-assay mean sum of squares from ANOVA, and these variances are fit to the variance model.
  - -Curve Residuals. The variance model is fit from the residuals of the individual dilution curves simultaneously from the pooled assays.
- The pooled response variances are then fit using an appropriate model equation.



The χ^2 Statistic ($RSSE_{nonparallel}$) behaves predictably



- The χ^2 statistic follows a predictable distribution if the variance is correctly modeled from a pool of historical assays, regardless of its magnitude or nature.
- -Parallelism thresholds can be established empirically.
- -Parallelism thresholds can be made lenient or strict, depending upon the needs and cost/benefit determinations.

Summary

- The Parallelism of Residuals (extra sum of squares) method for determining similarity of regressions is firmly established in the mathematical and bioassay literature, and widely used in many disciplines and industries.
- This classical statistical approach should be included in the USP <111> chapter as it is in the European Pharmacopeia 5.3.
- Comparing "linear regions" of nonlinear curves can fail to detect substantial dissimilarity between samples. Many bioassays do not have linear regions.
- The inability to determine accurate confidence regions for complex nonlinear curves is why the equivalence approach failed to correctly determine 10 of 20 expected results in this study.
- When correctly weighted with an adequate regression model, such as the 5PL for asymmetric dose-response curves, the extra sum of squares method yields reliable parallelism metrics, $RSSE_{nonpar}$ (χ^2 statistic) and χ^2 probability, that are accurate measures of the amount of parallelism between two curves, and solves the problems inherent in other approaches.
- The data from this study are available on www.brendan.com