

2011 MBSW Presentation

Statistical Perspectives Based on a Decade of Experience from Immunogenicity Cut Point Assessments

Wendell C. Smith, PhD
B2S Consulting™

wendells@b2s-stats.com | www.b2s-stats.com

Acknowledgements

- Ron Bowsher, B2S Consulting™
- Rocco Brunelle, B2S Consulting™
- Viswanath Devanarayan, Abbott Laboratories
- AAPS-sponsored white papers
 - *Mire-Sluis et al., 2004, JIM (design elements for ADA immunoassays)*
 - Koren et al., 2007, JIM (*immunogenicity testing strategy*)
 - *Shankar et al., 2008, JPBA (validation of ADA immunoassays)*
 - Gupta et al., 2011, JPBA (NAb assay validation)
- Regulatory Guidelines
 - EMEA guideline on immunogenicity assessment of biotechnology-derived therapeutics, 2007
 - FDA draft guidance, 2009

Presentation Outline

- Immunogenicity background information
 - Definition and concerns
 - Assay methods and test strategy
- Cut point study design
 - Design factors
 - Sample allocation
- Screening cut point data analysis methods
 - AAPS white paper recommendations
 - B2S Consulting™ standard approach
 - Example 1: Direct ELISA
 - Example 2: Bridging ECL
- Cut point application issues

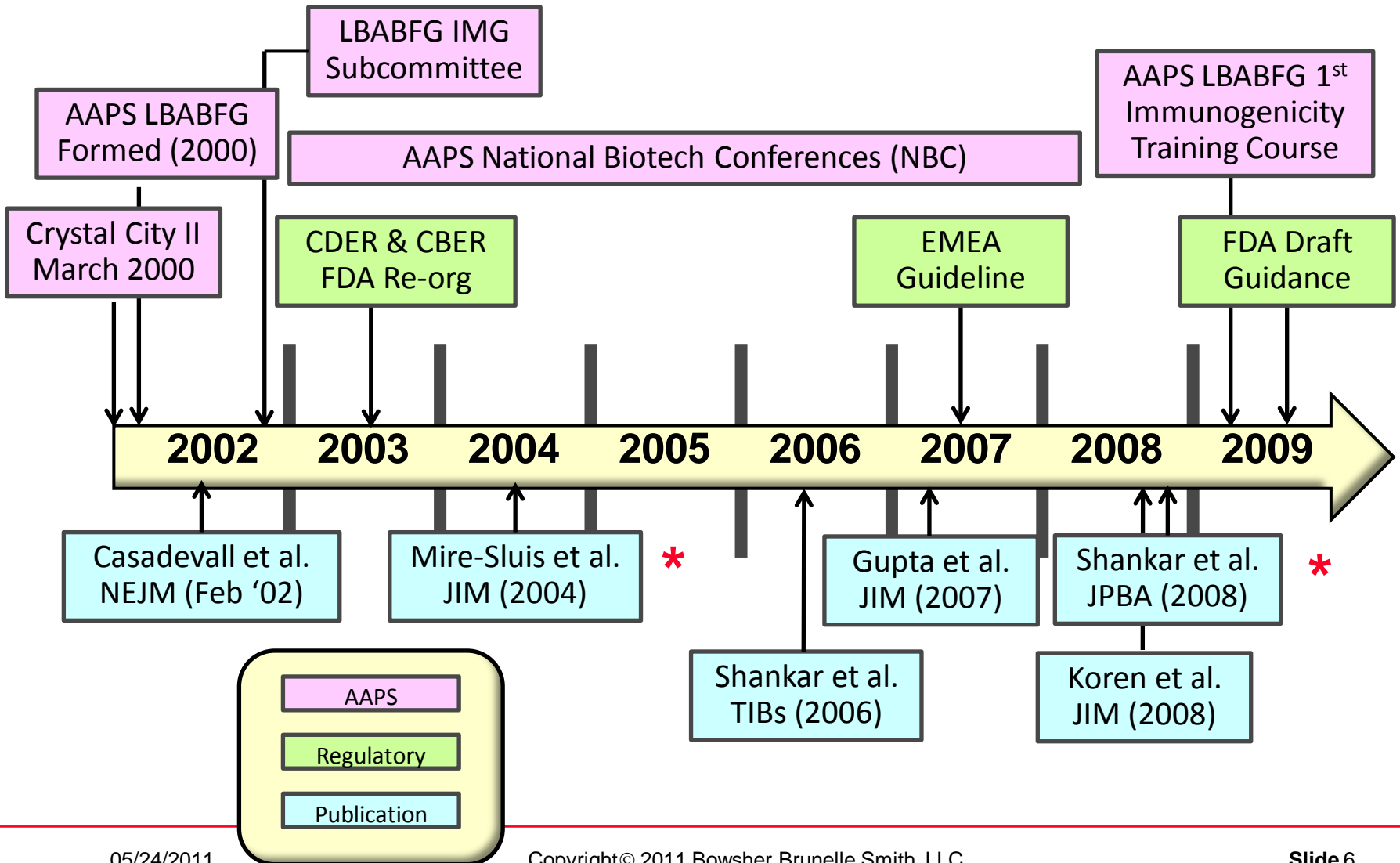
What is immunogenicity?

- Ability to provoke an immune response
- Two (2) types of immunogenicity:
 - “Unwanted” (formation of anti-drug antibodies, ADAs)
 - “Wanted” (formation of antibodies against vaccine antigens)
- Unwanted immunogenicity is the major safety concern for biotech drugs
- We can not reliably predict ADA incidence or severity of adverse drug reactions (ADR)
- Regulatory agencies highly recommend use of a risk-based strategy to evaluate “unwanted” immunogenicity
 - Risk = Likelihood of developing Abs x Consequence of Ab development
 - Detection and characterization of ADA is a key component of all strategies

ADA Regulatory Concerns

Concern	Clinical Outcome
Safety	<ul style="list-style-type: none">▪ ADA causes hypersensitivity reactions▪ ADA neutralize activity of an endogenous equivalent resulting in deficiency syndrome
Efficacy (PD)	<ul style="list-style-type: none">▪ ↑ or ↓ in efficacy resulting from a change in biotherapeutic half-life or biodistribution
PK	<ul style="list-style-type: none">▪ Altered PK caused by ADA results in a change in dosage level
None	<ul style="list-style-type: none">▪ Despite ADA generation, there are no discernable clinical effects /sequelae

Biotech's Immunogenicity Decade (2000's)



Immunogenicity Testing – 1st White Paper



**Numerous statistical design/method
recommendations**

Journal of Immunological Methods 289 (2004) 1–16



Standardization

Recommendations for the design and optimization of
immunoassays used in the detection of host antibodies against
biotechnology products

Anthony R. Mire-Sluis^{a,*}, Yu Chen Barrett^b, Viswanath Devanarayan^c,
Eugen Koren^d, Hank Liu^e, Mauricio Maia^f, Thomas Parish^g, George Scott^h,
Gopi Shankarⁱ, Elizabeth Shores^j, Steven J. Swanson^d, Gary Taniguchi^{k,†},
Daniel Wierda^l, Linda A. Zuckerman^m

Immunogenicity Testing – 2nd White Paper

Journal of Pharmaceutical and Biomedical Analysis 48 (2008) 1267–1281



Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba



Review

Recommendations for the validation of immunoassays used for detection of host antibodies against biotechnology products

Gopi Shankar^a, Viswanath Devanarayan^b, Lakshmi Amaravadi^c, Yu Chen Barrett^d, Ronald Bowsher^e, Deborah Finco-Kent^f, Michele Fiscella^g, Boris Gorovits^h, Susan Kirschner^{i,1}, Michael Moxness^j, Thomas Parish^k, Valerie Quarmby^l, Holly Smith^m, Wendell Smithⁿ, Linda A. Zuckerman^o, Eugen Koren^{p,*}

^a Clinical Pharmacology Sciences, Centocor Research & Development Inc., Radnor, PA 19087, USA

^b Global Exploratory Statistics, Abbott Laboratories, Parsippany, NJ 07054, USA

^c Preclinical & Clinical Development Sciences, Biogen Idec, Cambridge, MA 02142, USA

^d Clinical Discovery, Bristol-Myers Squibb Company, Princeton, NJ 08543, USA

^e Millipore, BioPharma Services Division, 15 Research Park Dr., St. Charles, MO 63304, USA and B2S Consulting, Beech Grove, IN 46107, USA

^f Drug Safety Research and Development, Pfizer, Groton, CT 06340, USA

^g Human Genome Sciences, 14200 Shady Grove Road, Rockville, MD 20850, USA

^h Drug Safety and Metabolism, Wyeth, Pearl River, NY 10965, USA

ⁱ Office of Biotechnology Products, CDER, Food and Drug Administration, Bethesda, MD 20892, USA

FDA Immunogenicity Draft Guidance

- Scope – ADA to Therapeutic Proteins
- Focus – Clinical investigation
- Guidance for assays:
 - ADA detection
 - ADA confirmation
 - Neutralizing Abs
- Also relevant for the evaluation of immune data from preclinical studies
 - Not predictive of man
 - Interpretation of Tox / pharm data
 - May reveal potential Ab related tox
- FDA supports evolving assay approach

Guidance for Industry Assay Development for Immunogenicity Testing of Therapeutic Proteins

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact (CDER) Susan Kirshner at 301-827-1731, or (CBER) Office of Communication, Outreach, and Development at 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

December 2009
CMC

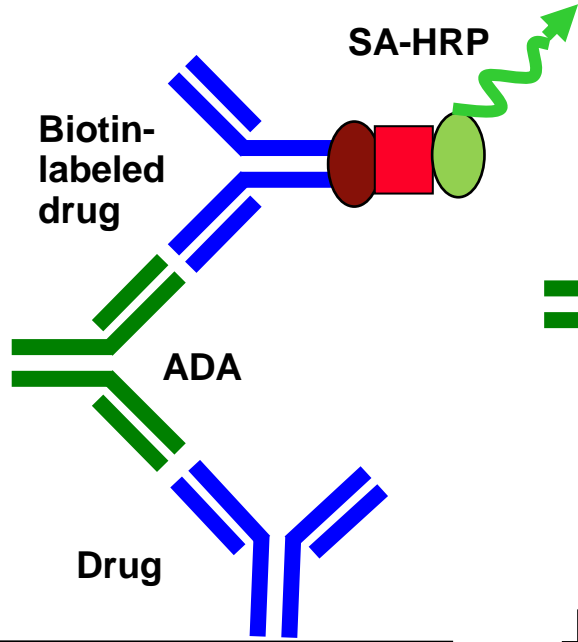
Biotechnology Protein Drugs / Assay Designs

Drug Type	Example	Assay Design
Monoclonal Antibody	Infliximab (Remicade), Avastin (Bevacizumab)	Bridging ELISA / ECL
Therapeutic protein	Erythropoetin (Epogen) hGH (Humatrope)	RIA, ELISA, ECL
Peptide	Insulin (Humulin), PTH ₁₋₃₄ (Forteo')	RIA, ELISA
Fusion Protein	Etanercept (Enbrel)	ELISA, ECL

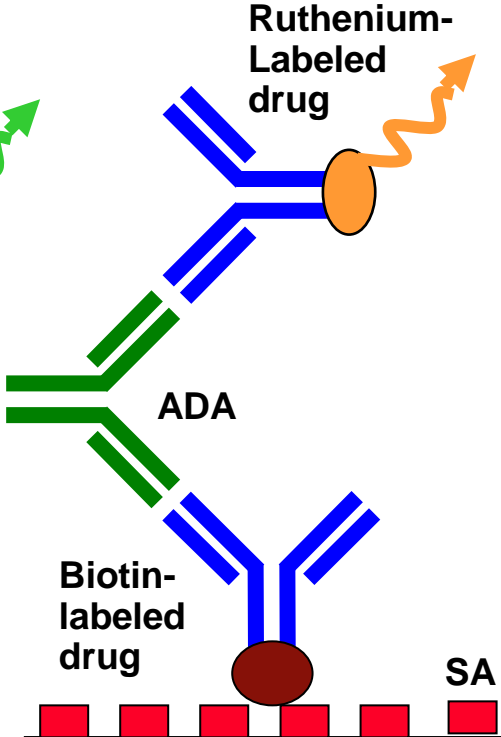
Assay Designs	Analytical Response
ELISA: Enzyme linked immunosorbant assay	Absorbance (OD)
ECL: Electrochemiluminescence	RLU / ECL
RIA / RIPA: Radioimmunoassay	%B/T

Assay Design Formats (Monoclonal Ab)

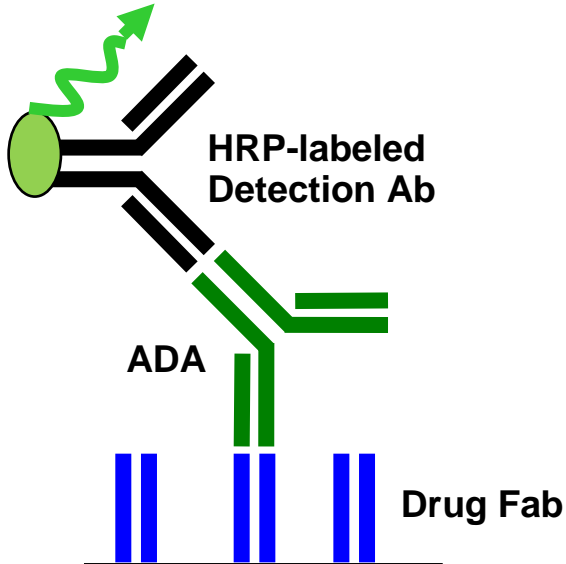
Bridging ELISA



Bridging ECL

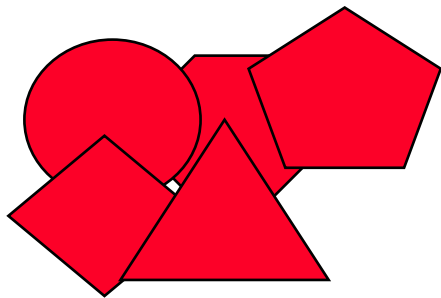


Direct ELISA

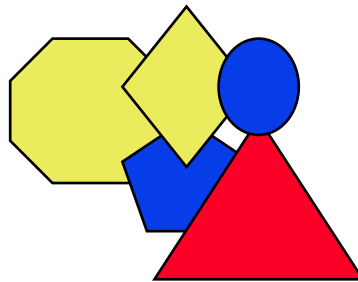


'Uncertainty Principle' of Anti-Drug Antibody Validations¹

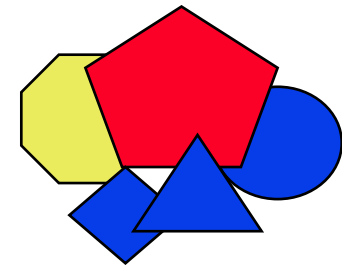
- Every sample has distinct mixture of isotypes, affinities, avidities (valency), epitope specificities, antibodies conc.
- Every sample is likely to differ in these characteristics from every other sample, including the positive control
- In normal bioanalysis practice, it is unknown how the characteristics differ from sample-to-sample



Positive Control



Sample 1



Sample 2

¹ Courtesy of Bonita Rup, Pfizer

For Immunogenicity applications...

- Immunogenicity assays are considered to be Quasi-Quantitative → sample result is reported in continuous units of a sample property (i.e., assay signal)

Why?

- Reference standards do not exist to reflect the Ab affinities and proportions in patient samples.
- Due to the lack of similarity between standard and test samples, use of a calibration curve to report assay results will likely introduce additional error in the identification and quantification of Ab+ samples.

ADA Four-tiered Test Strategy

- Tier 1: Identify “reactive” samples
 - Samples with signal greater than screening cut point
- Tier 2: Identify “Ab+” samples by testing reactive samples in the absence and presence of drug
 - Samples with percent inhibition greater than confirmatory cut point
- Tier 3: Determine a sample titer value by serial dilution of Ab+ samples in Tier 2
 - Titer is based on the screening cut point, and the value can be either continuous (requires interpolation) or discrete
- Tier 4: Evaluate neutralizing effects of antibodies
 - Based on cell-based bioassay using Ab+ samples

Cut Point Definitions

- Screening cut point: Level of **assay signal** at or above which a sample is defined to be putative positive ('reactive') and below which it is defined to be negative.
 - Determined statistically from the level of binding in drug-naïve samples
 - Binding may be nonspecific (due to assay background and sample matrix components) or specific (due to pre-existing endogenous or anti-drug antibody)
- Confirmatory cut point: Level of **signal inhibition** at or above which a (reactive) sample is judged to have specific anti-drug antibody
 - Determined by testing (reactive) drug-naïve samples in the absence and presence of drug

Cut Point Study Design

Cut Point Study: Sample Lots

- ≥ 50 individual drug-naïve normal human serum (**NHS**) sample lots
 - Usually purchased commercially
 - Usually but not always derived from a single individual
 - Individuals are assumed to be normal healthy adults and/or having a specific disease state (i.e. - diabetic)... no clinical history
 - Assumed to be drug-naive and antibody negative
- Negative base pool (**NBP**) sample lot
 - Created by pooling individual lots (after screening)
 - NBP usually becomes the assay negative control (NC)
 - Need sufficient volume of NBP to support in-study sample analyses
- Low, mid and high positive control lots (**LPC**, MPC, HPC)
 - Prepared by spiking the NBP lot with surrogate Ab

Cut Point Study: Design Example

Lot Group	Lot ID	Analyst 1			Analyst 2		
		Run 1 07-Oct-10	Run 2 08-Oct-10	Run 3 11-Oct-10	Run 4 07-Oct-10	Run 5 08-Oct-10	Run 6 11-Oct-10
A (N=17)	L01 – L17	Plate 1	Plate 3	Plate 2	Plate 1	Plate 3	Plate 2
B (N=17)	L18 – L34	Plate 2	Plate 1	Plate 3	Plate 2	Plate 1	Plate 3
C (N=17)	L35 – L51	Plate 3	Plate 2	Plate 1	Plate 3	Plate 2	Plate 1

Study Design Features

- Each NHS lot is tested once in each of ≥ 6 assay runs (3 per analyst).
- For testing across assay runs, lots are grouped into 'k' equal-size subgroups where k is the number of plates in a run.
- Lots in a subgroup are tested together on a single plate of each run.
- Across runs, each subgroup is tested an equal number of times on the k ordered assay plates.
 - Latin Square Design

ELISA Cut Point Study: Design Factors

Systematic (Fixed) Effects

- Subject disease state: NHA, T2D, RA, ...
- Sample lot (assay) group: A, B, C
- Assay analyst: AAA, BBB
- Assay plate order: P_1, P_2, P_3

Biological factors

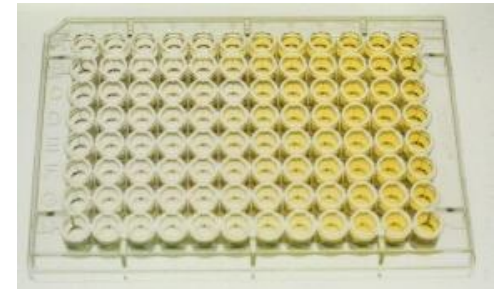


Random Effects

- Subject sample lot: L_1, \dots, L_{51}
- Assay run: $R_1, R_2, R_3, \dots, R_6$ (3 per analyst)
- Assay plate: $N=18$ (3 per run)
- Residual

ELISA Plate Design

- Sample assay result
 - Mean optical density (OD) from 2 wells (adjacent columns)
 - Result is accepted if $CV \leq 25\%$ for OD values from 2 wells ??
- Number of test results (with drug absent)
 - NHS: $N = 1$ per lot for subgroup (17)
 - NBP: $N \geq 2$ (i.e., 1 each at front, middle and back of plate)
 - LPC: $N = 2$ (i.e., 1 each at two split plate locations)
- Maximum number of NHS lots per plate is less when lots are also tested with drug



96-well ELISA
microtiter plate
8 rows, 12 cols)

Screening Cut Point Data Analysis

Mire-Sluis Illustration (1 run)

A.R. Mire-Sluis et al. / Journal of Immunological Methods 289 (2004) 1–16

9

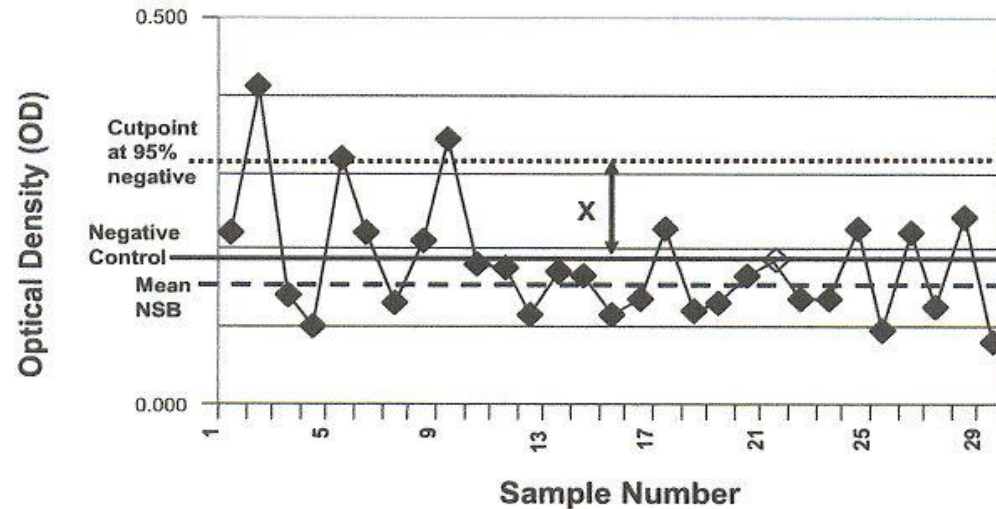


Fig. 2. An illustration of the selection of an assay cut point based on the nonspecific binding (NSB) of negative control matrix samples. Thirty negative controls are assayed (closed diamonds) and a mean NSB optical density (OD) is established (dashed line). An upper 95% confidence limit above the NSB mean is estimated (dotted line) that should result in 5% false positives. In order to account inter-assay variability, a negative control is selected to be incorporated in every assay (open diamond). A normalization factor (X) is derived from the difference in OD of the negative control and the OD of the 95% limit. After each assay, the normalization factor is added to the OD of the negative control to set the 95% cut point of the assay.

Need to log-transform OD values???

Data Analysis Topics (White papers)

- Target false positive error rate of 5% for screening CP [versus 0.1% (or 1%) for confirmatory CP]
- Determine appropriate data transformation
- Remove samples (lots) with preexisting specific anti-drug antibodies.
- Remove statistical “outliers” resulting from non-specific matrix factors [How about analytical factors?]
- Confirm distributional assumptions
 - Normality
 - Variance homogeneity
- Cut point determinations
 - Fixed versus floating
 - Parametric versus nonparametric

B2S Consulting™

Standard Data Analysis Approach

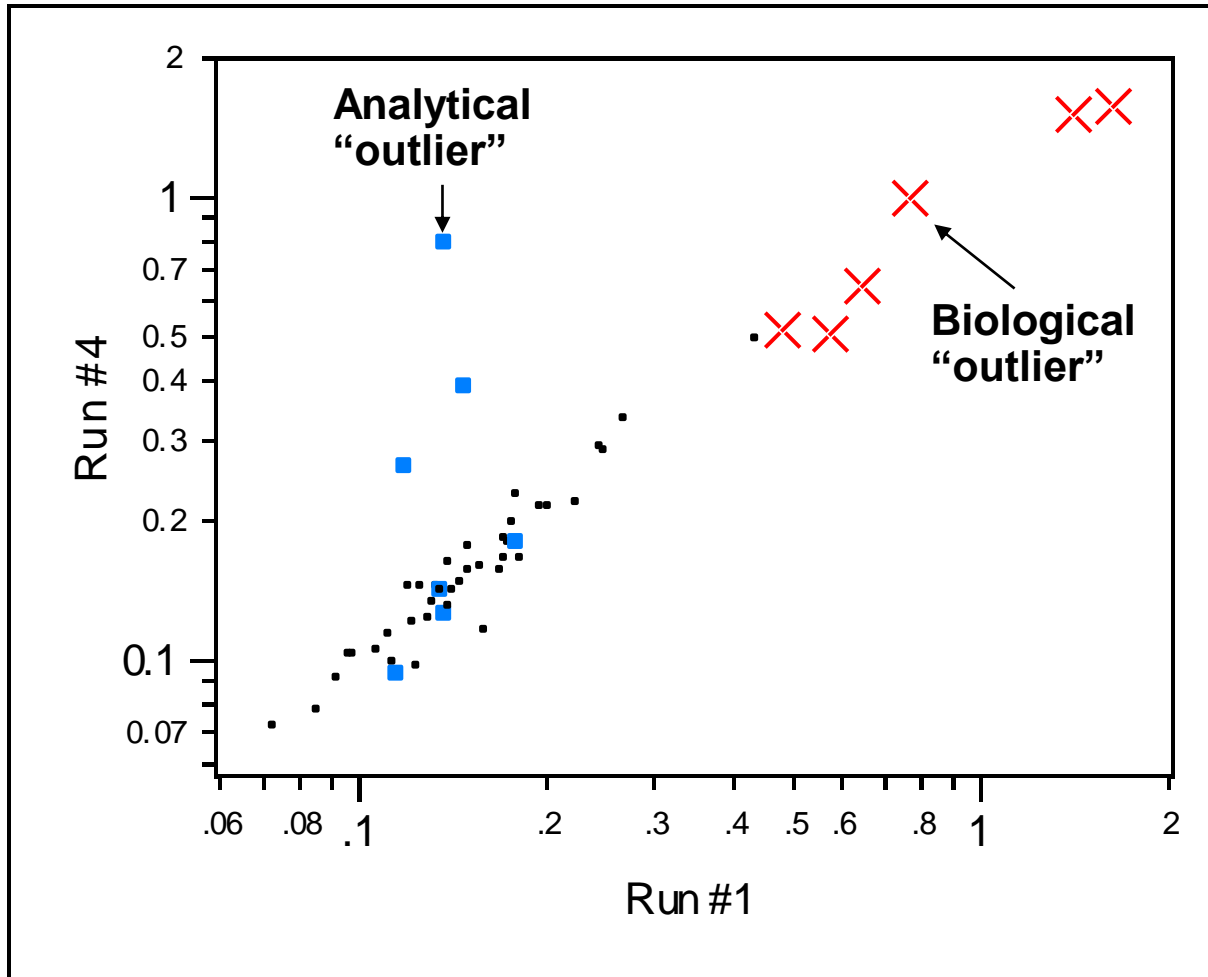
- Default data transformation is specified based on the assay design and experience from previous studies
- Linear mixed effects analysis of variance (ANOVA) model is specified to investigate sources of variation
 - Removal of biological and analytical statistical outliers
 - Assessment of fixed and random effects
 - Confirmation of ANOVA distributional assumptions
- Cut point estimation (fixed and floating)
 - Robust parametric procedure (Tukey biweight)
 - Nonparametric procedure (empirical 95th percentile)
- Cut point application
 - Evaluate benefit of floating cut point, based on the NBP mean value for a plate (or run), versus fixed cut point

Statistical Outliers

- Outliers are identified by the “outlier box-plot” criteria
 - Value $> Q3 + 1.5*(Q3-Q1)$ or $< Q1 - 1.5*(Q3-Q1)$
 - » $Q3 = 75^{\text{th}}$ percentile, $Q1 = 25^{\text{th}}$ percentile ($Q2 = \text{median}$)
- **Analytical outlier** is identified by applying criteria to ANOVA conditional residual values
- **Biological outlier** is identified by applying criteria to sample lot ANOVA best linear unbiased predictor (BLUP) values

Note: Outliers are excluded sequentially (1 at a time?) due to masking and/or lack of independence.

Outlier Illustration (2006 ELISA)



Note: In this example, biological outlier signals are not inhibited by drug. Mean OD for LPC is ~ 0.316.

Screening Cut Point Data Analysis

Example 1: **Direct ELISA**

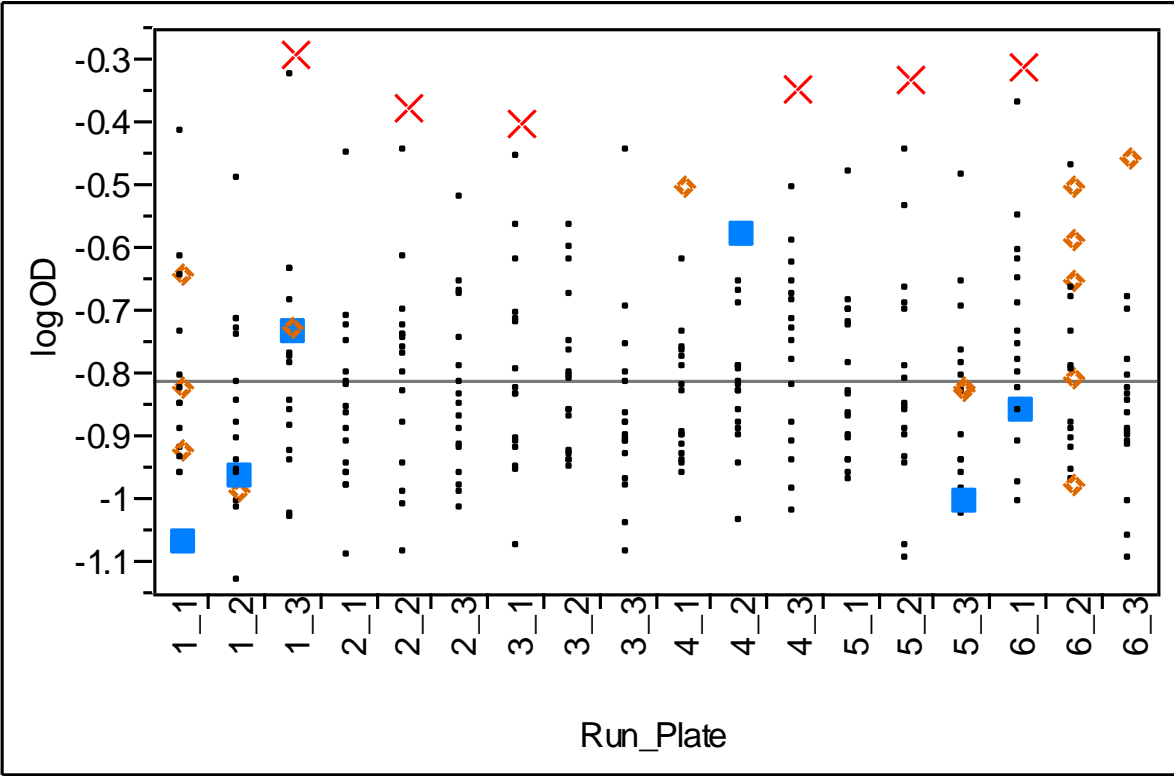
Example 1: Study Design

Lot Group	Lot IDs	Analyst 1			Analyst 2		
		Run 1 07-Oct-10	Run 2 08-Oct-10	Run 3 11-Oct-10	Run 4 07-Oct-10	Run 5 08-Oct-10	Run 6 11-Oct-10
A (N=17)	L01 – L17	Plate 1	Plate 3	Plate 2	Plate 1	Plate 3	Plate 2
B (N=17)	L18 – L34	Plate 2	Plate 1	Plate 3	Plate 2	Plate 1	Plate 3
(N=17)	L35 – L51	Plate 3	Plate 2	Plate 1	Plate 3	Plate 2	Plate 1

Example 1: Data-related Comments

- Total of 306 optical density (OD) values
 - 51 NHS sample lots (from drug-naïve NHA)
 - 6 assay runs
- Fourteen (14) values were excluded because $CV > 25\%$ for OD from duplicate wells
- Twelve (12) values were excluded as statistical outliers based on the linear mixed effects ANOVA of log (base 10) transformed OD values
 - 6 individual values identified as analytical outliers
 - 6 values from one lot identified as a biological outlier

Example 1: Data Scatterplot

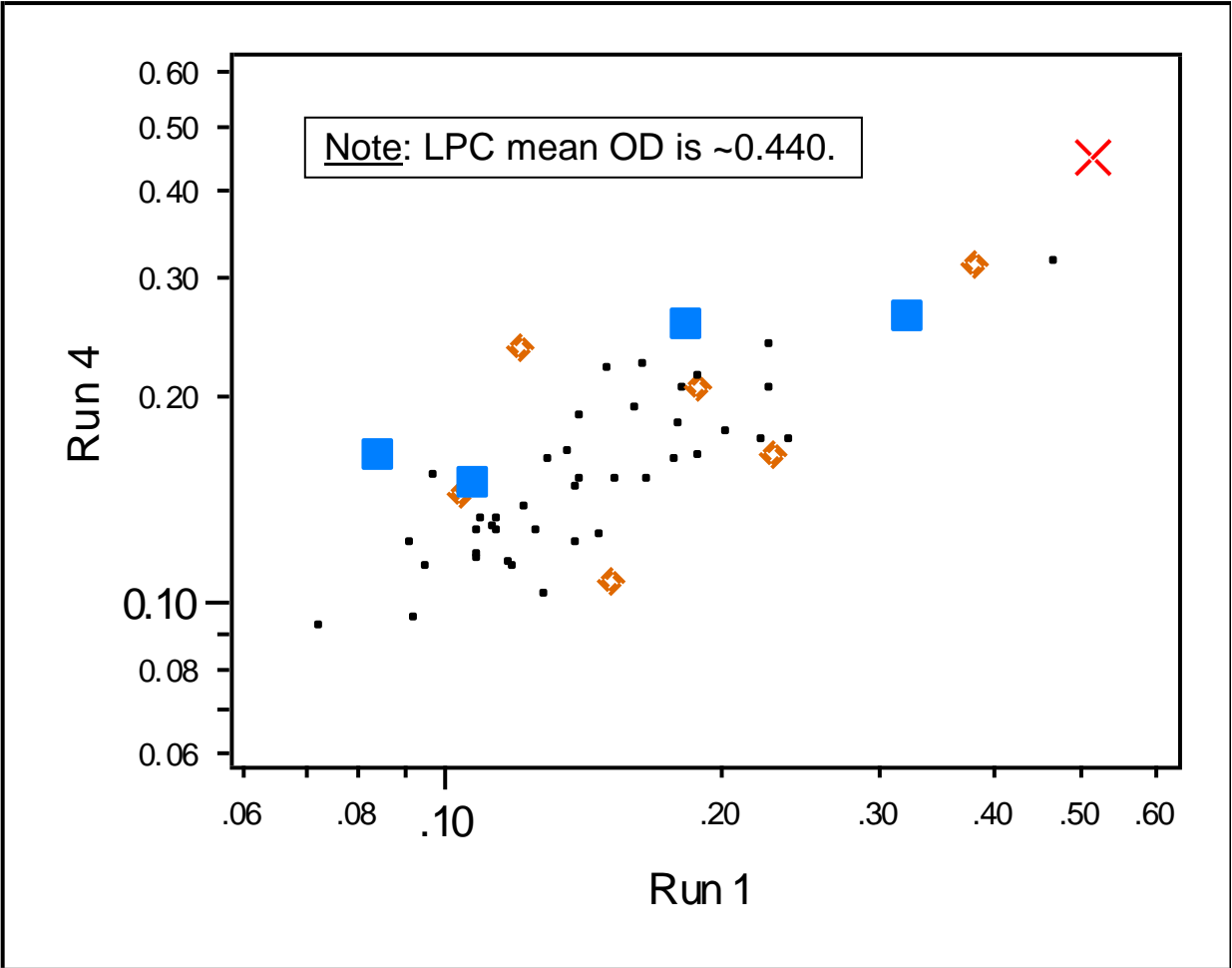


Biological outlier

Analytical outlier

CV > 25%

Example 1: Run 4 versus Run 1



Biological
outliers

Analytical
outlier

CV > 25%

Example 1: ANOVA Random Effects

Random Effect	Variance Estimate	Variance Ratio	Percent Of Total
Lot(Group)	0.020226	10.31	88.87
Run(Analyst)	0.000000	0.00	0.00
Assay Plate	0.000573	0.29	2.51
Residual	0.001961	1.00	8.62
Total	0.022760	-	100.0

Example 1: ANOVA Fixed Effects

Fixed Effect	Num DF	Den DF	P-value
Group	2	52.2	0.541
Analyst	1	10.0	0.236
Plate Order	2	10.0	0.562
Analyst*Plate Order	2	10.0	0.816

Diagnostic tests:

- > Normality of BLUPs (Shapiro-Wilk p-value = 0.093)
- > Normality of conditional residuals (S-W p-value = 0.137)
- > Intra-plate variance homogeneity (Levene p-value = 0.830)

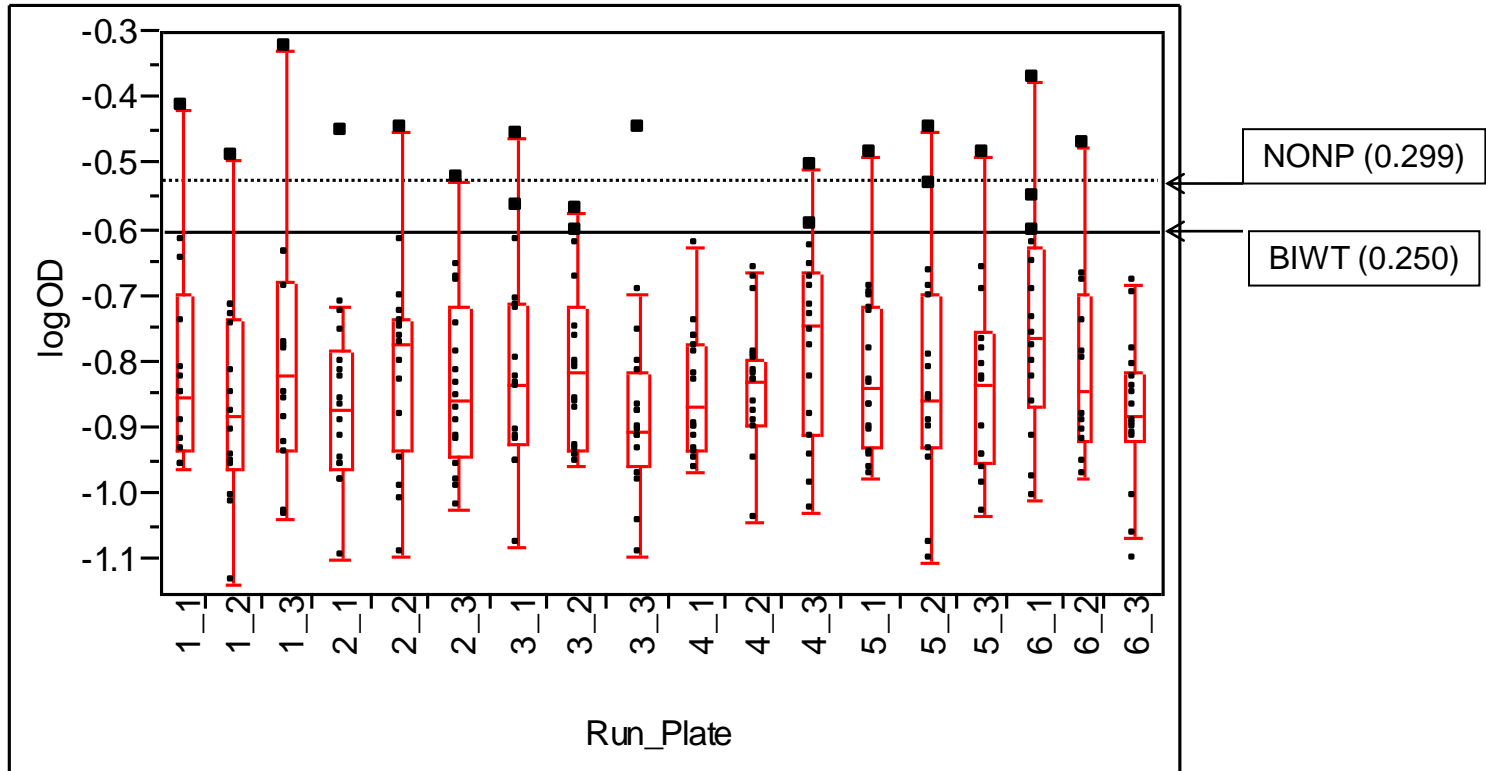
Example 1: Cut Point Estimates

Statistical Method		Data Level	Fixed (OD)	Floating (Ratio ¹)
Parametric	Classical	All plates	0.261	2.21
	Biweight	All plates	0.250	2.10
		Run level ²	0.251	2.13
Nonparametric	Biweight	All plates	0.299	2.54
		Run level ²	0.311	2.61

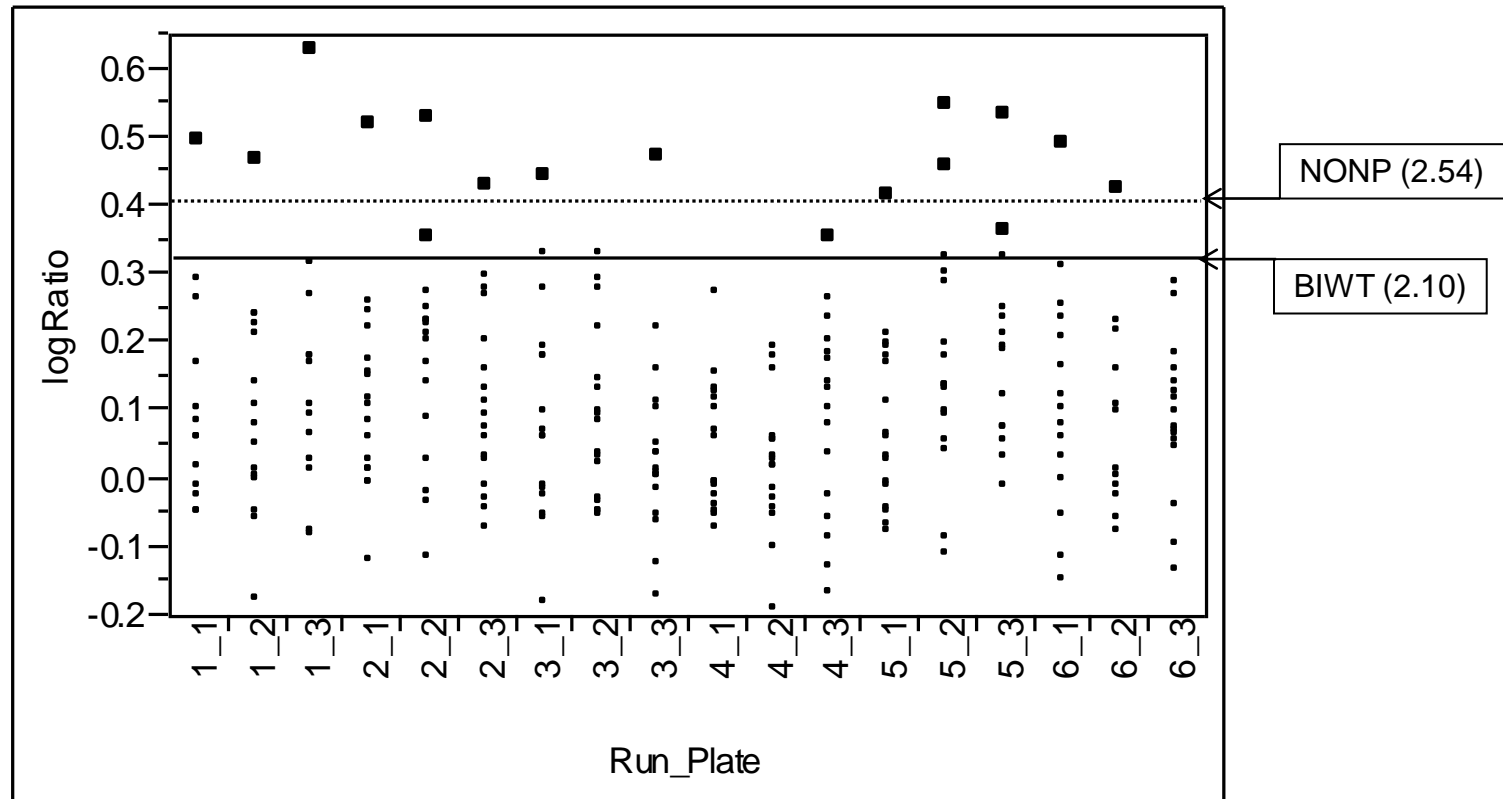
¹ Ratio is calculated by dividing each NHS OD value by the NBP geometric mean for the plate (or run → for run level estimates)

² Cut point is determined for each run and then pooled to obtain overall value (Shankar white paper)

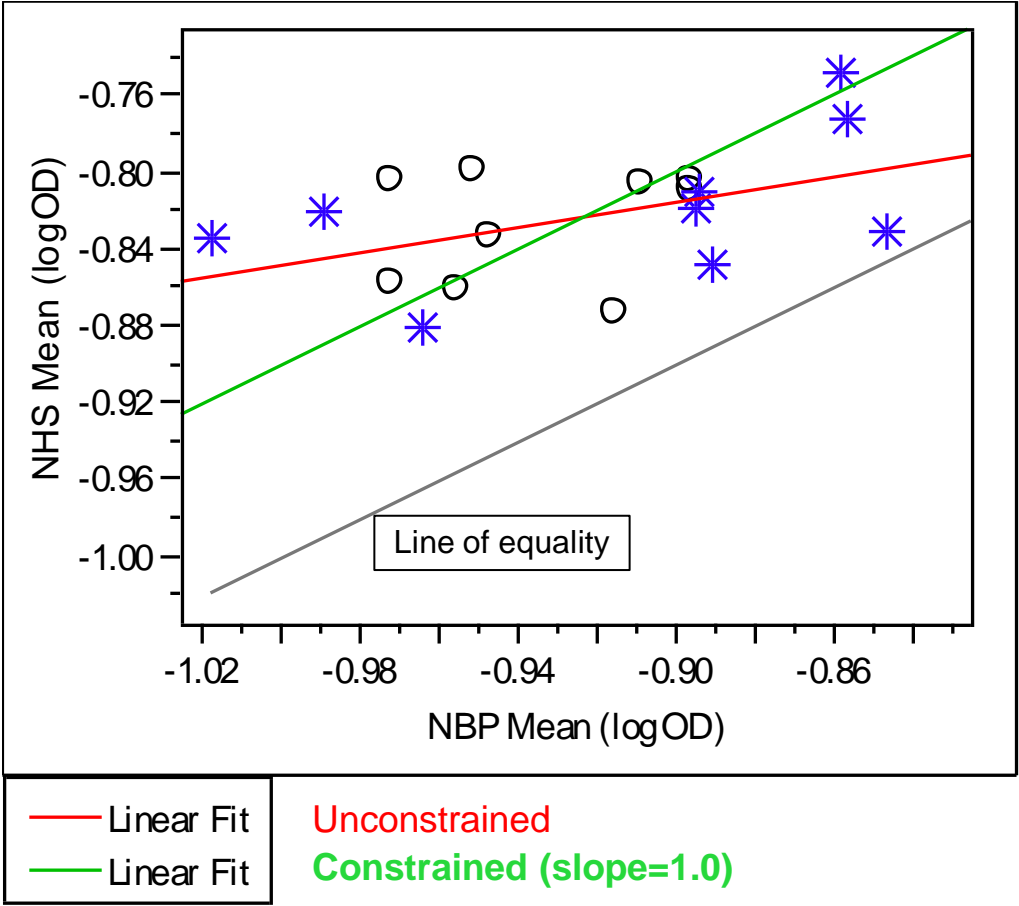
Example 1: Fixed Cut Point



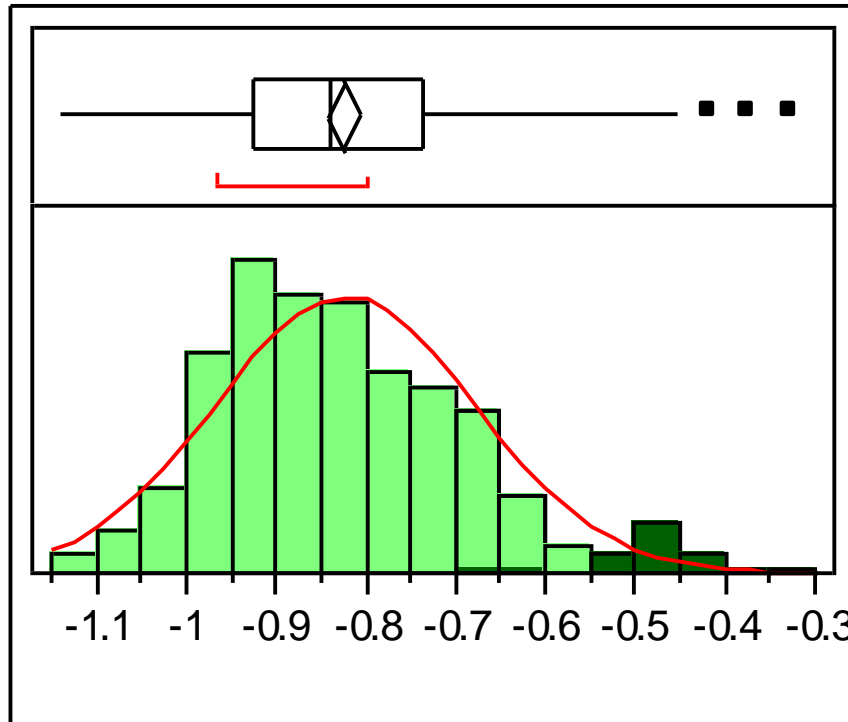
Example 1: Floating Cut Point Factor



Example 1: Plate mean values



Example 1: Histogram of log OD Values



Fitted Normal

Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.968963	<.0001 *

Note: Normality p-value is 0.087 if values from 3 higher lots are excluded (lots are biological outliers in 2-factor random effects model).

Sources of Non- normality

- Inappropriate data transformation
 - log-transformation generally works well
- Presence of a few samples with relatively high signals that are not excluded as outliers (previous slide)
- Significant difference between mean signal values among levels of an analytical fixed effect factor (i.e., analyst, plate order, ...)
 - Mean difference is often explained by NBP → floating cut point
- Significant difference between mean signal values among levels of a biological fixed effect factor (i.e., disease state, gender,...)
 - Mean difference is not explained by NBP → consider separate cut points

Screening Cut Point Data Analysis

Example 2: **Bridging ECL**

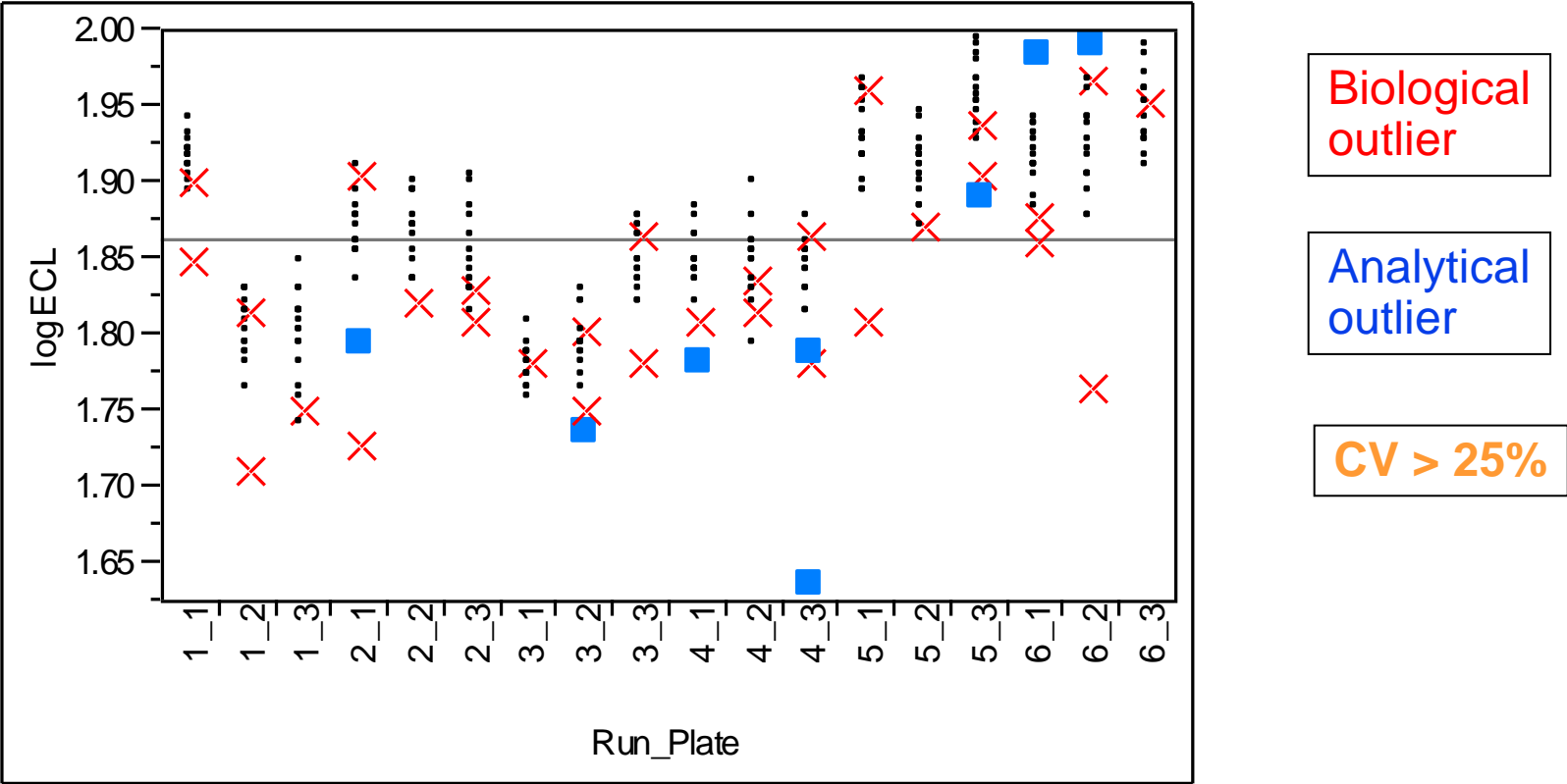
Example 2: Study Design

Lot Group	Lot IDs	Analyst 1			Analyst 2		
		Run 1 29-Jun-10	Run 2 30-Jun-10	Run 3 01-Jul-10	Run 4 29-Jun-10	Run 5 02-Jul-10	Run 6 01-Jul-10
A (N=17)	L01 – L17	Plate 1	Plate 3	Plate 2	Plate 2	Plate 3	Plate 1
B (N=17)	L18 – L34	Plate 2	Plate 1	Plate 3	Plate 3	Plate 1	Plate 2
C (N=17)	L35 – L51	Plate 3	Plate 2	Plate 1	Plate 1	Plate 2	Plate 3

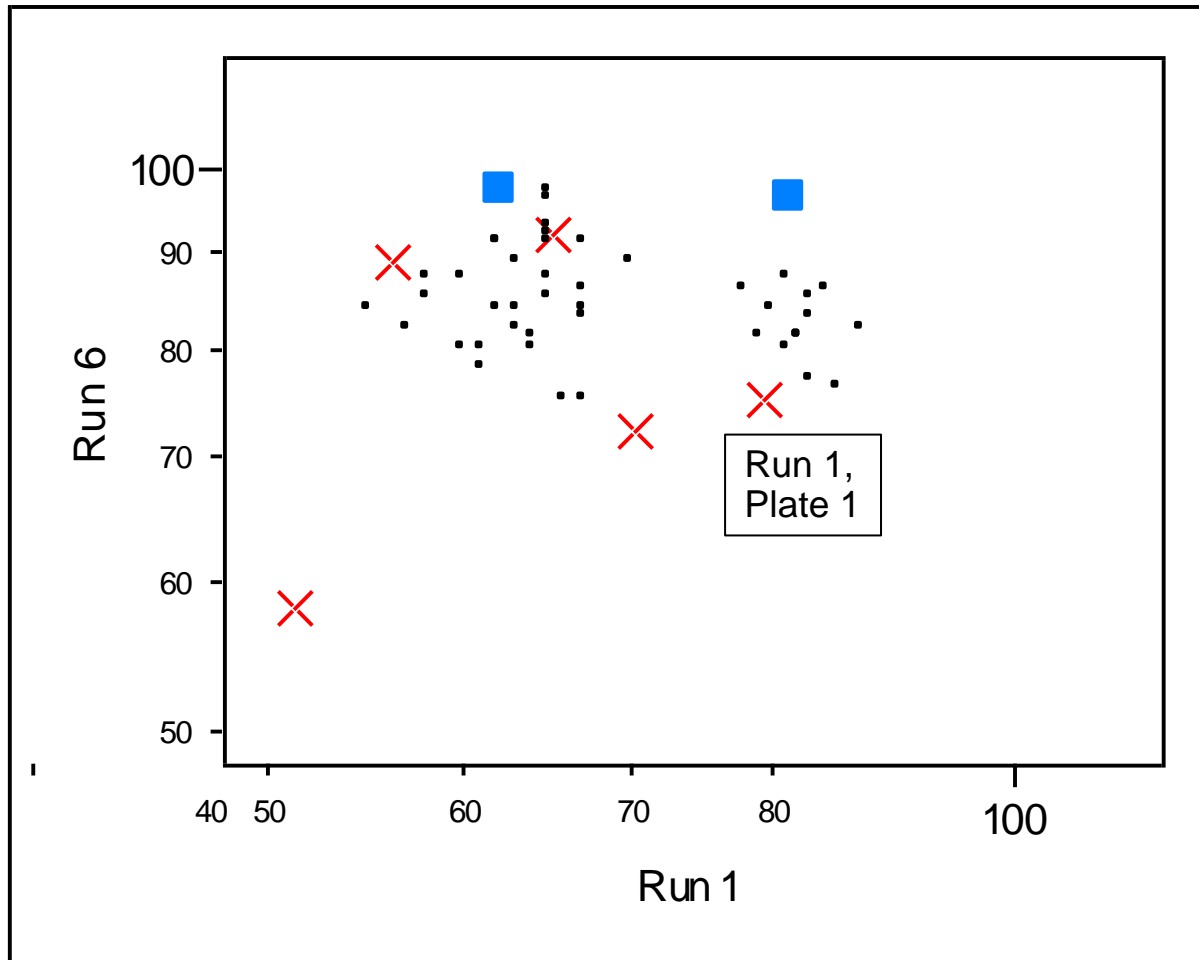
Example 2: Data-related Comments

- Total of 306 optical density (OD) values
 - 51 NHS sample lots (from drug-naïve NHA)
 - 6 assay runs
- Zero (0) values excluded because of high CV
- Thirty-eight (38) values excluded as statistical outliers based on linear mixed effects ANOVA of log (base 10) transformed OD values
 - 8 individual values identified as analytical outliers
 - 30 values from five (5) lots identified as biological outliers

Example 2: Data Scatterplot



Example 2: Run 6 versus Run 1



Biological
outlier

Analytical
outlier

CV > 25%

Example 2: ANOVA Random Effects

Random Effect	Variance Estimate	Variance Ratio	Percent Of Total
Lot(Group)	0.000030	0.07	1.00
Run(Analyst)	0.001040	2.42	34.64
Assay Plate	0.001501	3.49	50.02
Residual	0.000430	1.00	14.34
Total	0.003001	-	100.0

Example 2: ANOVA Fixed Effects

Fixed Effect	Num DF	Den DF	P-value
Group	2	6.1	0.641
Analyst	1	4.1	0.112
Plate Order	2	6.0	0.653
Analyst*Plate Order	2	6.0	0.689

Diagnostic tests:

- > Normality of BLUPs (Shapiro-Wilk p-value = 0.842)
- > Normality of conditional residuals (S-W p-value = 0.608)
- > Intra-plate variance homogeneity (Levene p-value = 0.033)←

p < 0.001
at run level

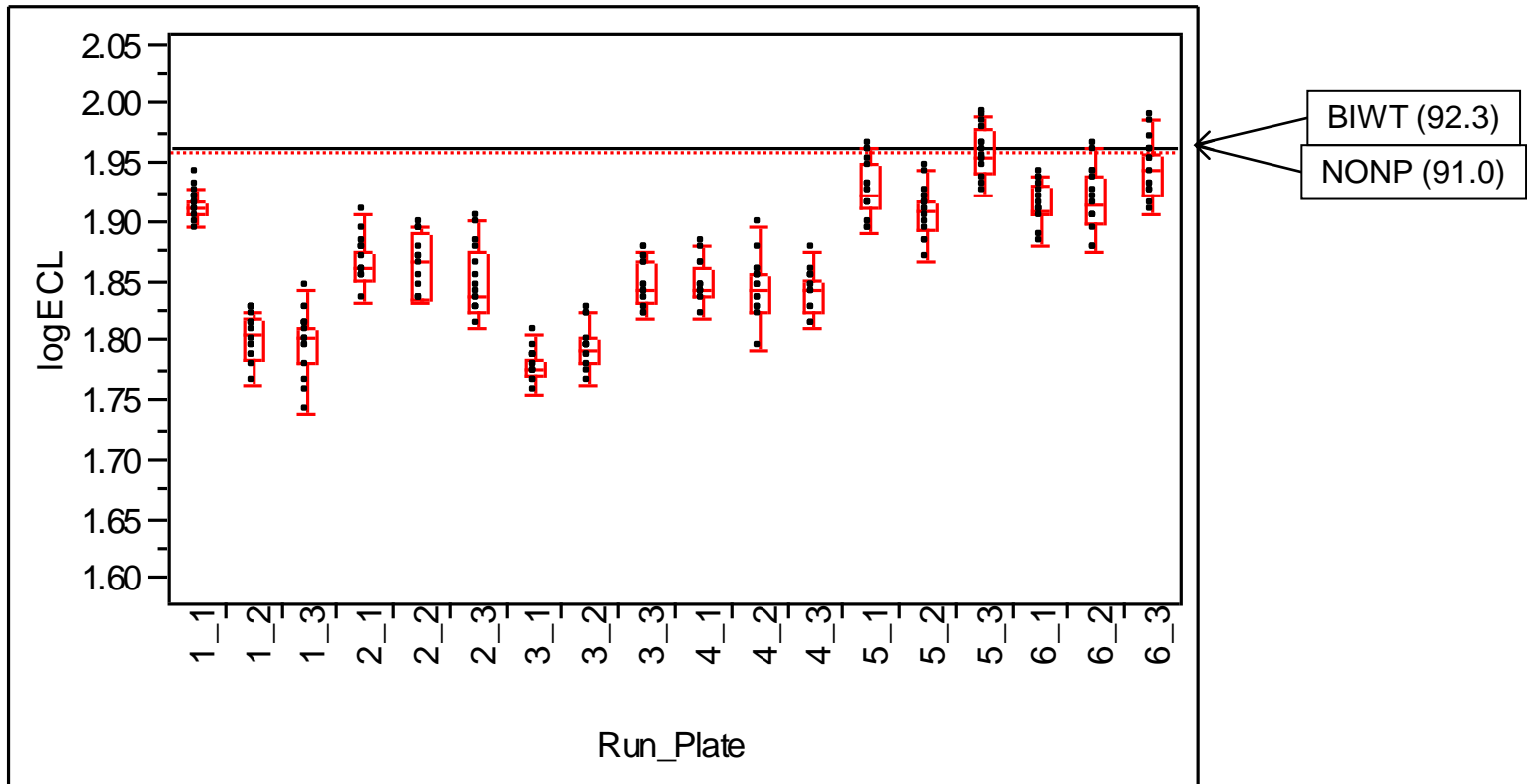
Example 2: Cut Point Estimates

Statistical Method		Data Level	Fixed (ECL)	Floating (Ratio ¹)
Parametric	Classical	All plates	91.7	1.15
	Biweight	All plates	92.3	1.14
		Run level ²	85.2	1.17
Nonparametric	Biweight	All plates	91.0	1.17
		Run level ²	84.1	1.17

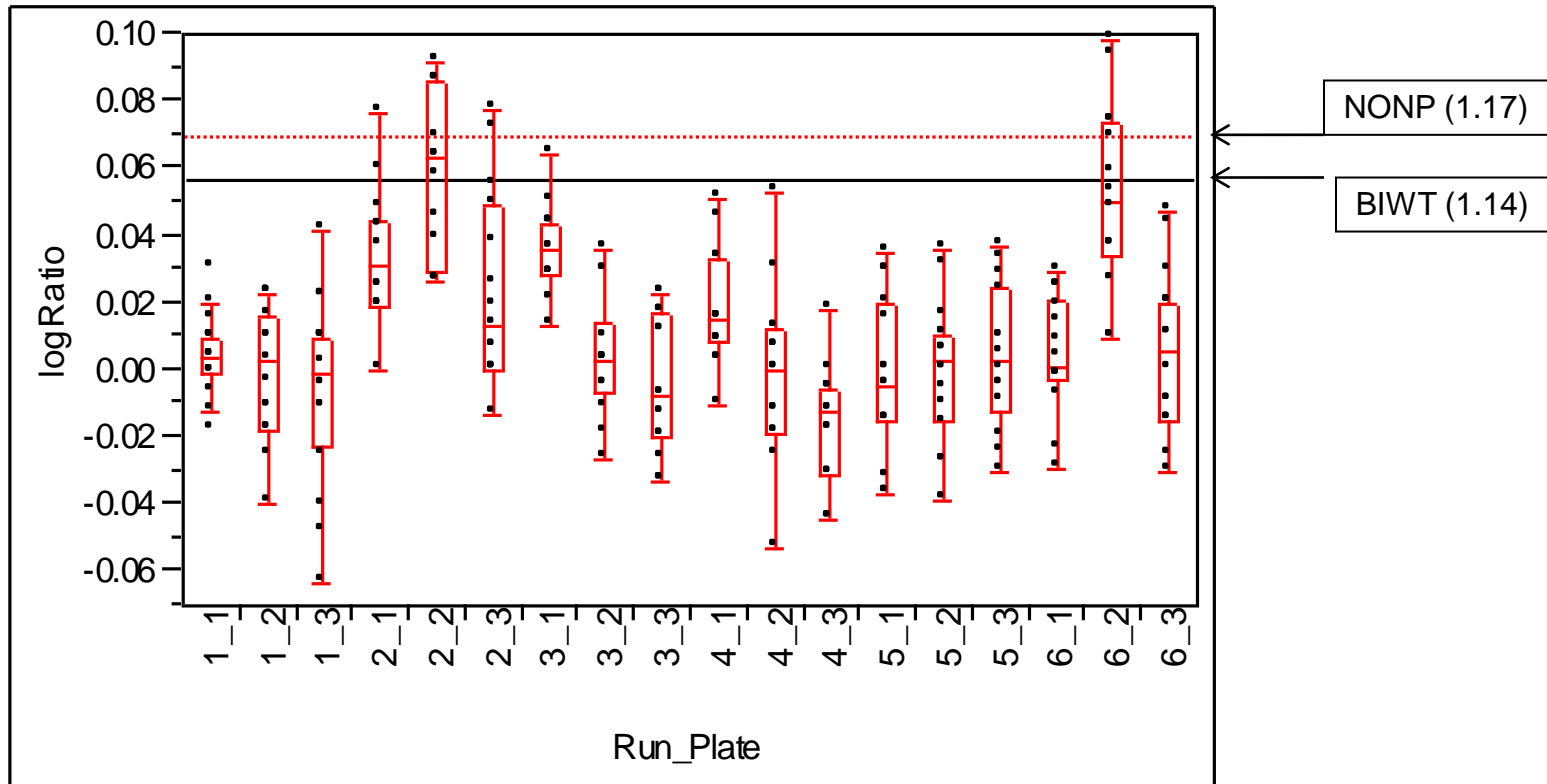
¹ Ratio is calculated by dividing each NHS ECL value by the NBP geometric mean for the plate (or run → for run level estimates)

² Cut point is determined for each run and then pooled to obtain overall value (Shankar white paper)

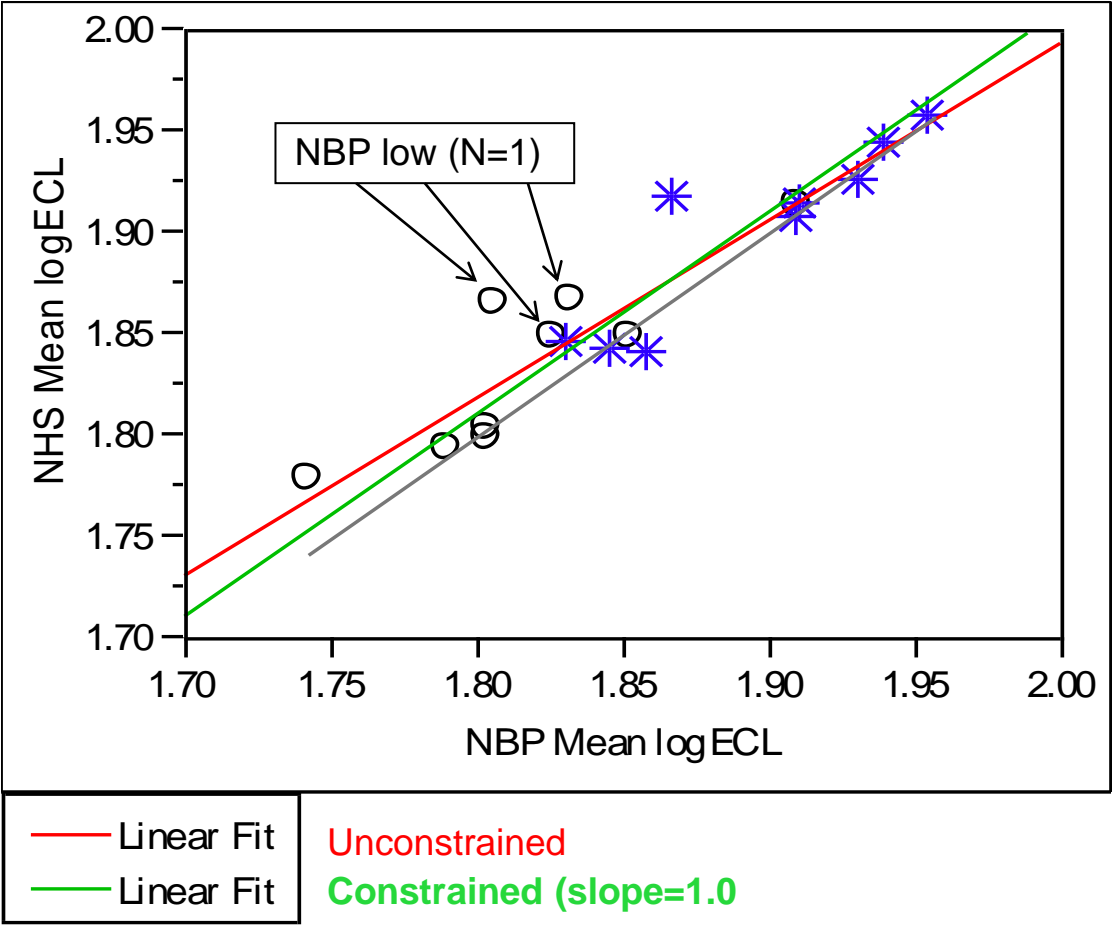
Example 2: Fixed Cut Point



Example 2: Floating Cut Point Factor

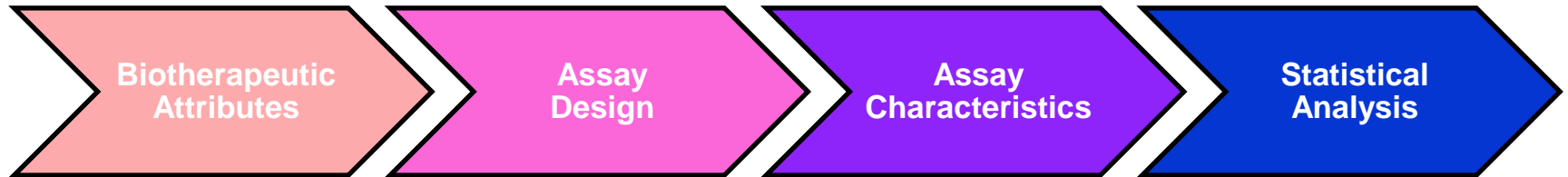


Example 2: Plate mean values



Summary / Conclusion

	2000	2011
Assay Design	ELISA	ECL
Assay Background	High	Low
Variance	High	Low
Sources of Variation	Biologic > Analytic	Analytic > Biologic



Advancement in biotherapeutics will lead to assay evolution which will drive progress in data-driven assignment of immunogenicity CPs

Cut Point Application

Application Issues

- Tier 1 screen (ADA Detection)
 - Is a cut point based on drug-naïve samples useful when the lot variance component is near 0? Should cut point be based on LPC?
 - Is a fixed cut point needed when laboratory scientists almost always prefer a (multiplicative) floating cut point?
 - How useful is a multiplicative floating cut point factor ≤ 1.0 \rightarrow negative control will be positive if equivalent to NBP
 - How should a cut point be determined and applied when the signal distribution for sample lots is bimodal (i.e., due to high percentage of samples with endogenous antibody present)?
 - Should the cut point be adjusted as a result of reagent changes (i.e., new conjugation lot)? If so, how?
- Tier 2 inhibition (ADA Confirmation)
 - Does it make sense to calculate a confirmatory cut point based on a 0.1% target false positive error rate when only 50 lots are tested?