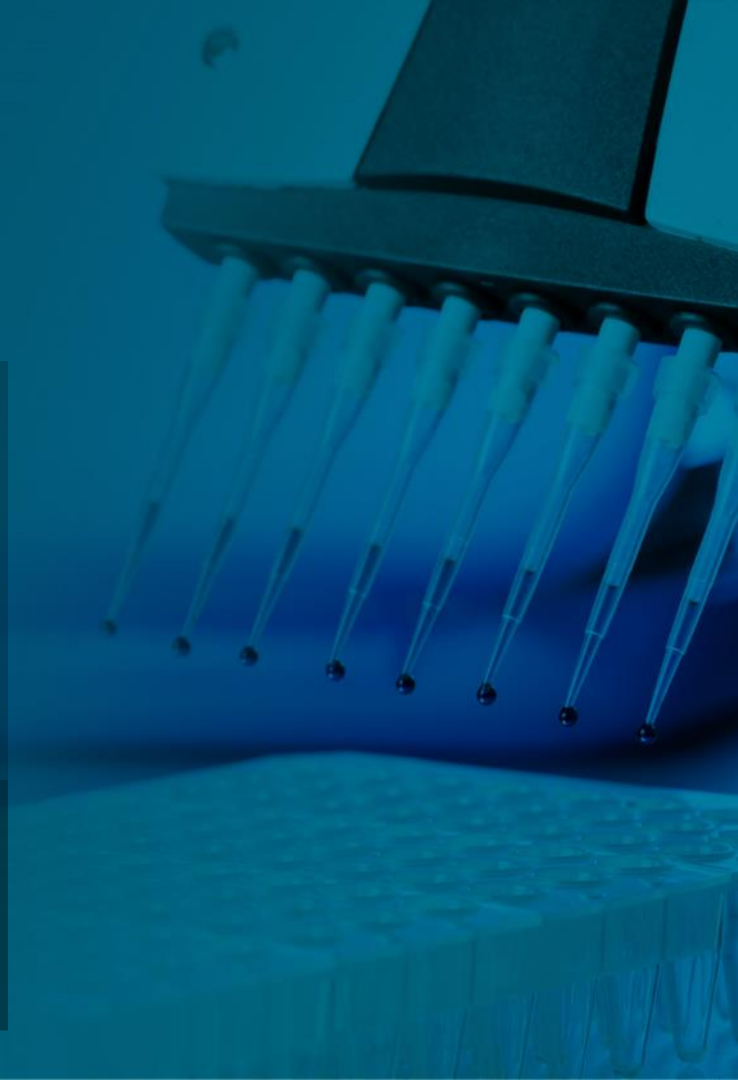
BioAgilytix

Considerations for Evaluation of Parallelism in Single and Multiplex Biomarker Ligand-Binding Assays

Afshin Safavi, Ph.D.
Founder and CSO, BioAgilytix

Todd Lester, PMP
Bioanalytical Project Manager, BioAgilytix

November 17th, 2016 – EBF Open Symposium

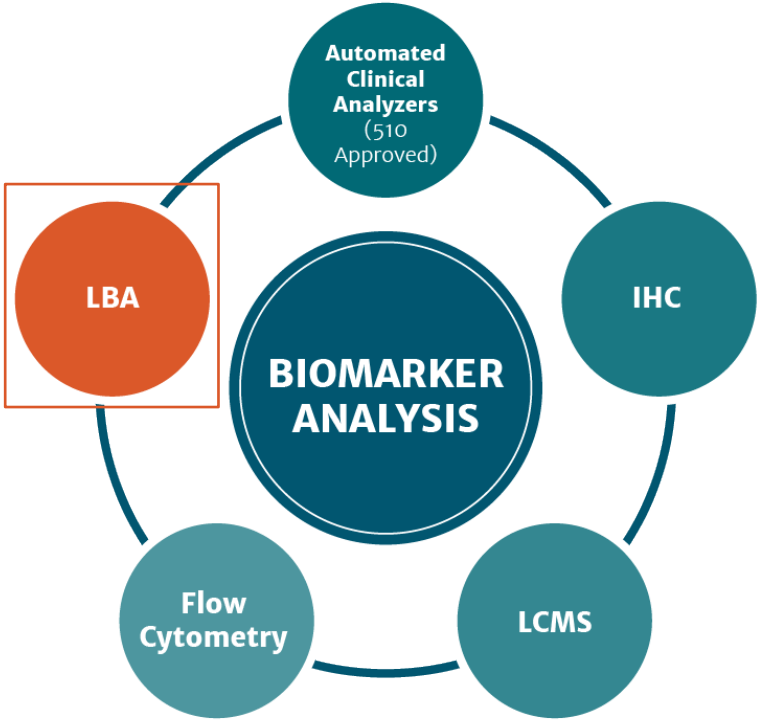


General Goals of Biomarker Use

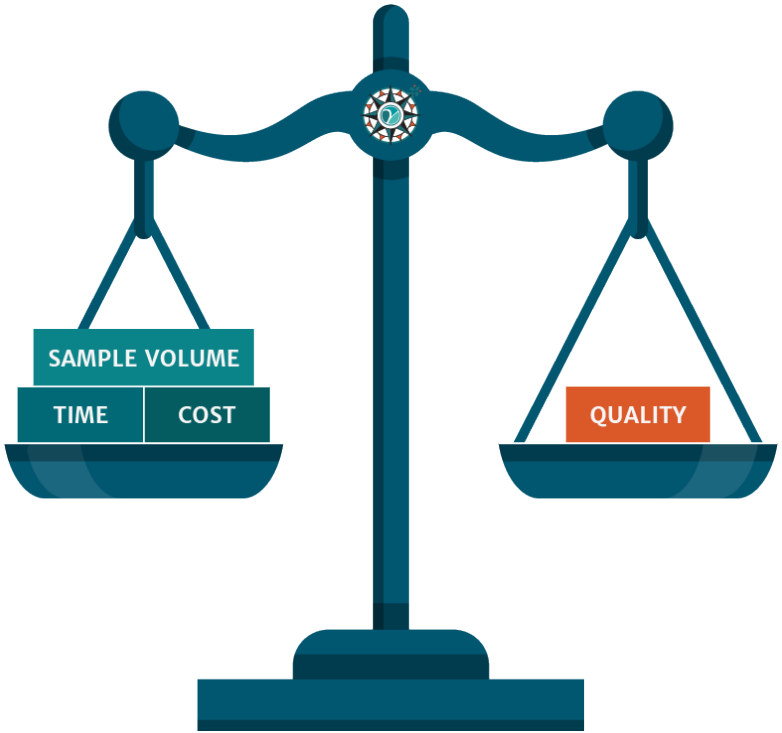
Drug development has entered a critical time dictated by the business needs to **produce more in less time and reduce cost.**

Companies are relying more on biomarkers to assess the efficacy, safety, and mechanism of action of drugs to get a **“go” or “no go” decision.**

Biomarker **Analysis**



Balancing Act



Quality Biomarker Data

Topics Covering Today

Relative Accuracy

Parallelism

Kit and Reagent Lot to Lot Variability

Let us discuss the concepts and review a few case studies

Key Publications on Biomarker Validation and Analysis in LBA

2001 FDA Guidance:
Although no reference to biomarkers and only PK, but still relevant.

2006 Whitepaper by Lee et al:
First key white paper (*Fit for Purpose*) with some consensus on biomarker validation in bioanalytical labs.

2011 EMEA Guidance:
Although no reference to biomarkers and only PK, but still relevant.

2012 Whitepaper by EBF:
European consensus paper.

2012 Whitepaper by GCC:
Global CRO consortium paper.

2013 Draft FDA Guidance:
In discussion-one page biomarker.

There is **no final guidance available** for biomarker analysis for LBAs

The World of **LBA Biomarkers…**

1. The concept behind ligand-binding assays is based on **measuring an analyte concentration via immunoreactivity of an antibody (or a binding partner) to the analyte of interest.**

2. For biomarker assays, the calibrators are typically either recombinant or purified materials and therefore **most often not identical to the endogenous form being measured.**

3. **Hard to find a "clean matrix"** to perform spike-recovery studies.

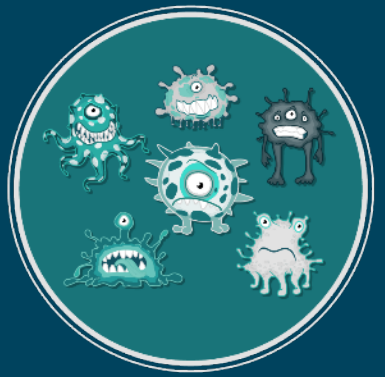
Final Goal: is the assay suitable to quantify the analyte of interest reliably and reproducibly?

Biomarkers are **NOT** PK Assays

PK

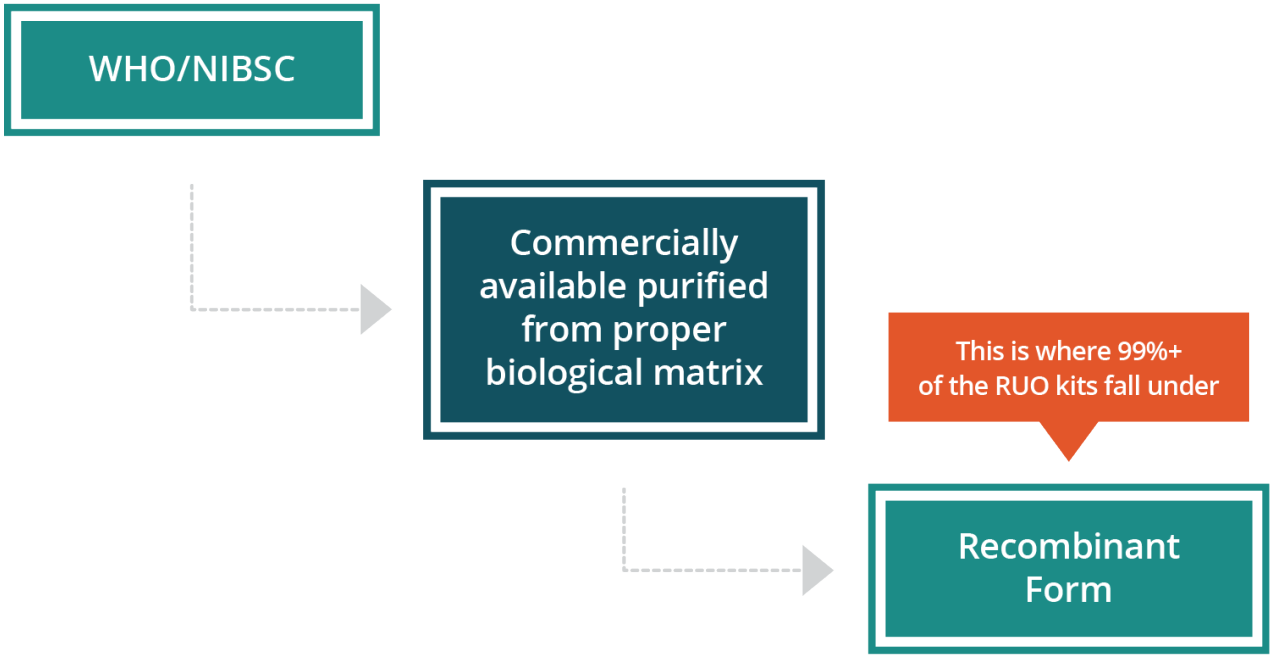
- ✓ Well characterized reference materials (CoA)
- ✓ Available in pure form
- ✓ Available in simple buffer/formulation

Biomarker

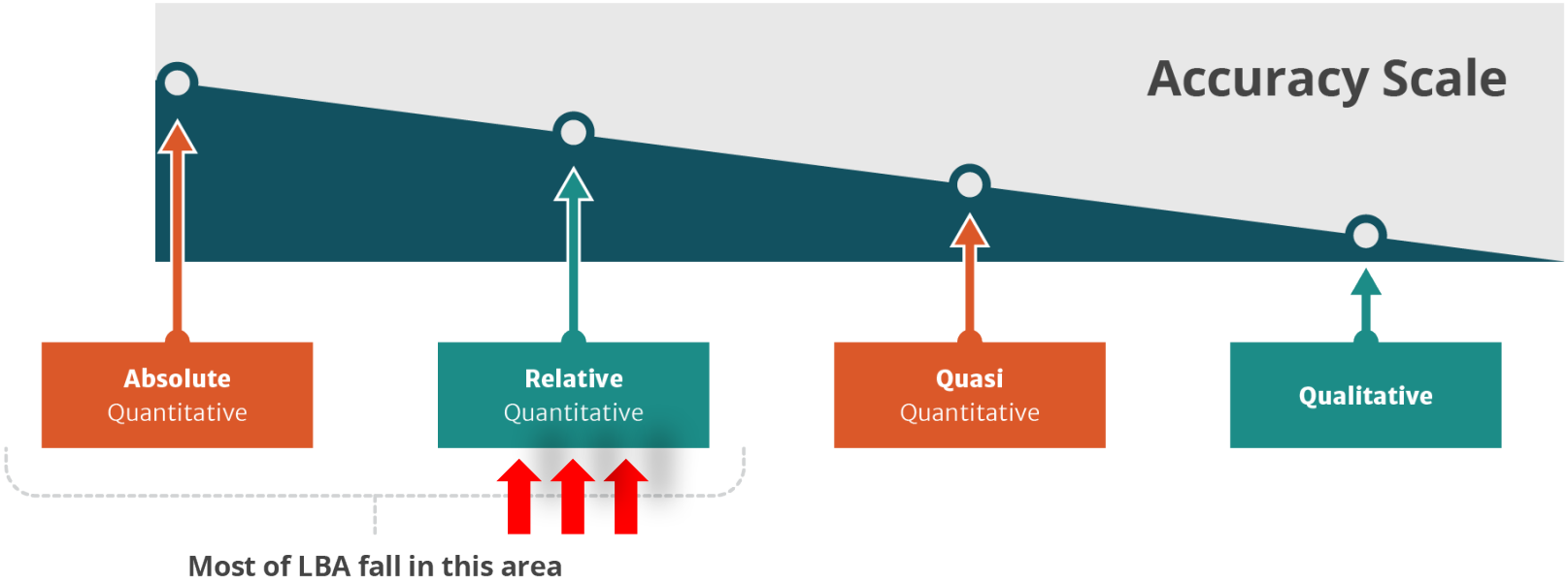


- ✓ Recombinant reference materials (Usually No CoA)
- ✓ Often not available in pure form
- ✓ Often Endogenous form available only in biological matrix

Reference Material/Calibrator Selection



Concept of Accuracy for Biomarker Analysis Using LBA ("Relative Accuracy" in Most Cases)



How Do We Use a RUO Kit to Support Our Clinical Biomarker Study?

Fit for Purpose Validation

- ✓ Standard curve precision and accuracy
- ✓ QC Precision and accuracy
- ✓ Calibration range
- ✓ Intra/Inter assay accuracy
- ✓ Intra/Inter assay precision
- ✓ Minimum required dilution
- ✓ **Parallelism**
- ✓ Specificity
- ✓ Selectivity
- ✓ Stability

What is Parallelism?

1. **A condition in which dilution of test samples does not result in biased measurements of the analyte concentration.**

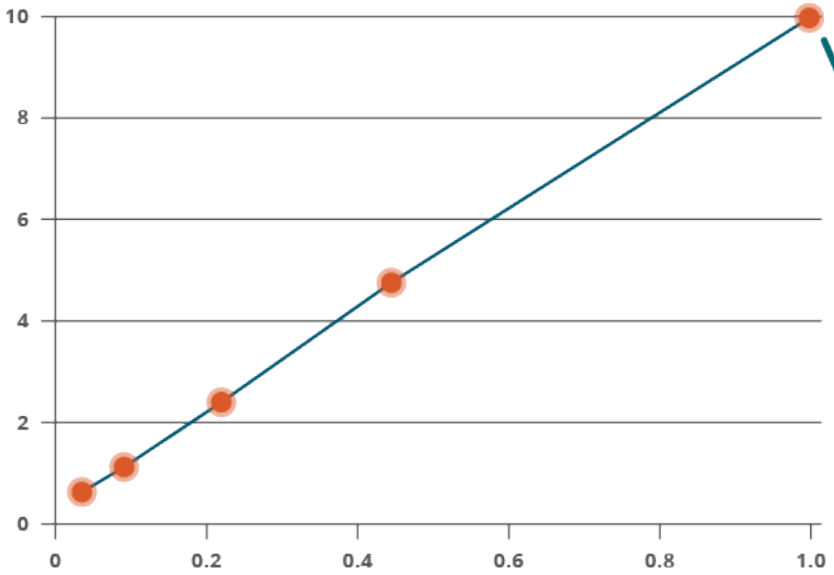
   "Thus, when a test sample is serially diluted to result in a set of samples having analyte concentrations that fall within the quantitative range of the assay, **there is no apparent trend toward increasing or decreasing estimates of analyte concentrations** over the range of dilutions."

   Miller et al., Pharm Research 18(9), 1373-1381, 2001

2. **Demonstration that the sample dilution response curve is parallel to the standard concentration response curve.**

   Lauren Stevenson and Shobha Purushothama, Bioanalysis 6(2), 185-198, 2015

How do **Parallelism and Dilutional Linearity** Differ?



Dilutional Linearity

Spike the matrix with drug (analyte) and then serially dilute.

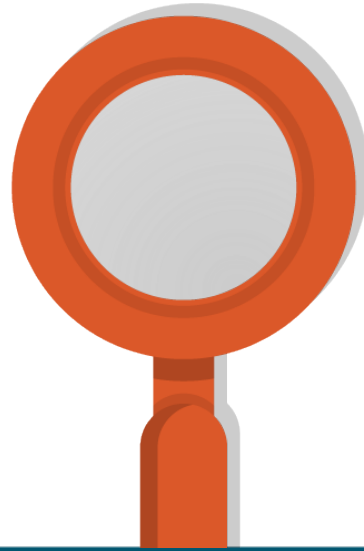
Parallelism

Find a sample with high endogenous level of analyte and then serially dilute.

How to Evaluate Parallelism

What is the general industry practice?

- ✓ **Screen and identify preferably at least 6 samples with a high level of the analyte.**
  - This practice varies from company to company, usually from 3 to 10 samples.
- ✓ **Perform serial dilutions (usually 2-fold) with the objective to obtain ≥3 dilutions falling within the assay range.**
  - This is very much assay and platform dependent as the dynamic range of the assay may vary.
  - Example: Getting 3 to 4 diluted points on an ELISA is typical but you may be able to get 6+ on an MSD or DELFIA.
- ✓ **Multiplex assays: identifying samples that have high levels of all biomarkers can be very challenging and are usually not available.**
  - In such cases, it may be necessary to use different samples for different biomarkers.
- ✓ **What if there are no samples with a high concentration of the biomarker(s)?**



What are the Acceptance Criteria for Biomarker Parallelism Assessment?

There is no clear requirement but more an industry consensus with regard to the acceptance criteria for parallelism:

1. CV of ≤ 30% amongst the in-range measurements back-calculated concentration to neat concentration (some labs are going with ≤ 25%).
   - *Another way to say it: Precision of the diluted samples should be < 30%.*
2. No trend is observed with increasing sample dilution (somewhat qualitative).
3. Please note that fewer than 10 papers attempting to address the topic of parallelism have been published in the last 5 years, so what the "acceptance criteria" should be is at its infancy and somewhat superficial.
4. One can also argue that even if there should be acceptance criteria set for parallelism assessment, it should not be "pass" or "fail" since most of the biomarker work falls under "Fit for Purpose". Ultimately, it is important have the information on parallelism, but what you do with it depends on the intended use of the assay.
5. Acceptance criteria stringency may be set as tighter or loser as long as the scientific rationale is justified and documented.

When in the Assay Development Process Should Parallelism be Investigated?

1 Development Stage

Unlike PK assays, for evaluation of the parallelism of biomarker assays, **there is no need to wait for incurred samples (study samples) to be available.** One can screen a series of disease-state and/or normal samples to find a few suitable samples for an initial evaluation of parallelism.

- ✓ **Limitations:** most likely the clinical demographic and number of F/T not known and/or available.
- ✓ Still the data will be valuable to determine early on if there may be some assay limitations and issues due to the sample matrix.
- ✓ It also provides some preliminary information regarding the MRD, assay selectivity and potential LLOQ.

2 Pre-study Validation Stage

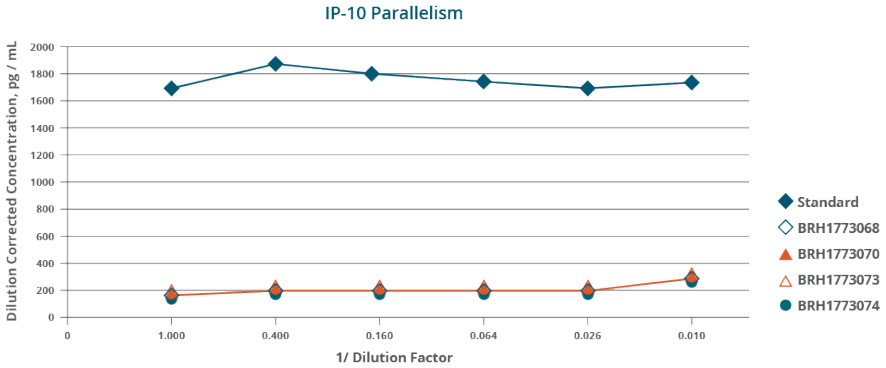
By this stage, **parallelism and/or any assay limitations should have been determined** and the main goal at this point would be final evaluation and documentation of any issues.

3 In-study Validation

Not required for exploratory biomarkers unless disease-state matrix was never tested or not available up to this stage. However, late stage biomarkers, biomarker studies done with an intention to develop diagnostics, and/or end-point biomarkers may require assessment of parallelism at this stage.

Case Study 1:

Typical Results for Parallelism



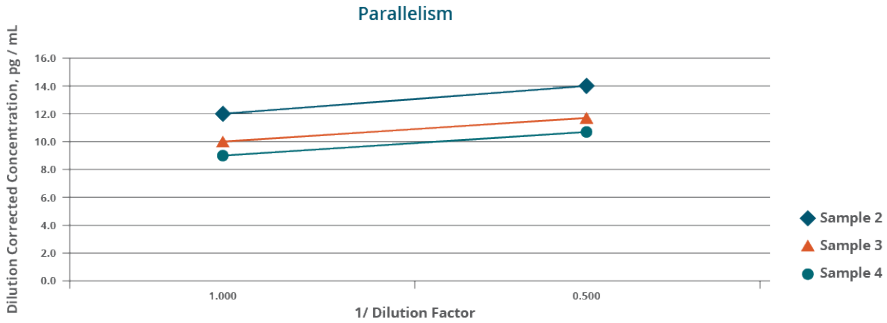
Dilution Factor		1P-10 Dilution Corrected Concentration (pg/mL)				
Dilution Fold	1/Dilution	Standard	BRH1773068	BRH1773070	BRH1773073	BRH1773074
1.00	1.000	1687	154	169	160	130
2.50	0.400	1892	176	232	217	193
6.25	0.160	1847	176	212	222	195
15.6	0.064	1725	176	226	231	196
39.1	0.026	1704	184	214	226	202
97.7	0.010	1704	212	241	257	231
Overall CV (%)		5.0	10.4	11.8	14.6	17.3

Case Study 2: Limited Parallelism Window

		Dilution Corrected bFGF Concentration (pg/mL)			
Dilution Fold	1/Dilution	Sample 1	Sample 2	Sample 3	Sample 4
1	1.000	BLQ	*4.03**	*2.95**	*160*
2	0.500	5.15	12.3	10.1	9.51
4	0.250	BLQ	13.7	11.4	10.9
8	0.125	BLQ	BLQ	BLQ	BLQ
16	0.0625	BLQ	BLQ	BLQ	BLQ
	CV (%)	**N/A**	**7.7**	**8.8**	**9.8**

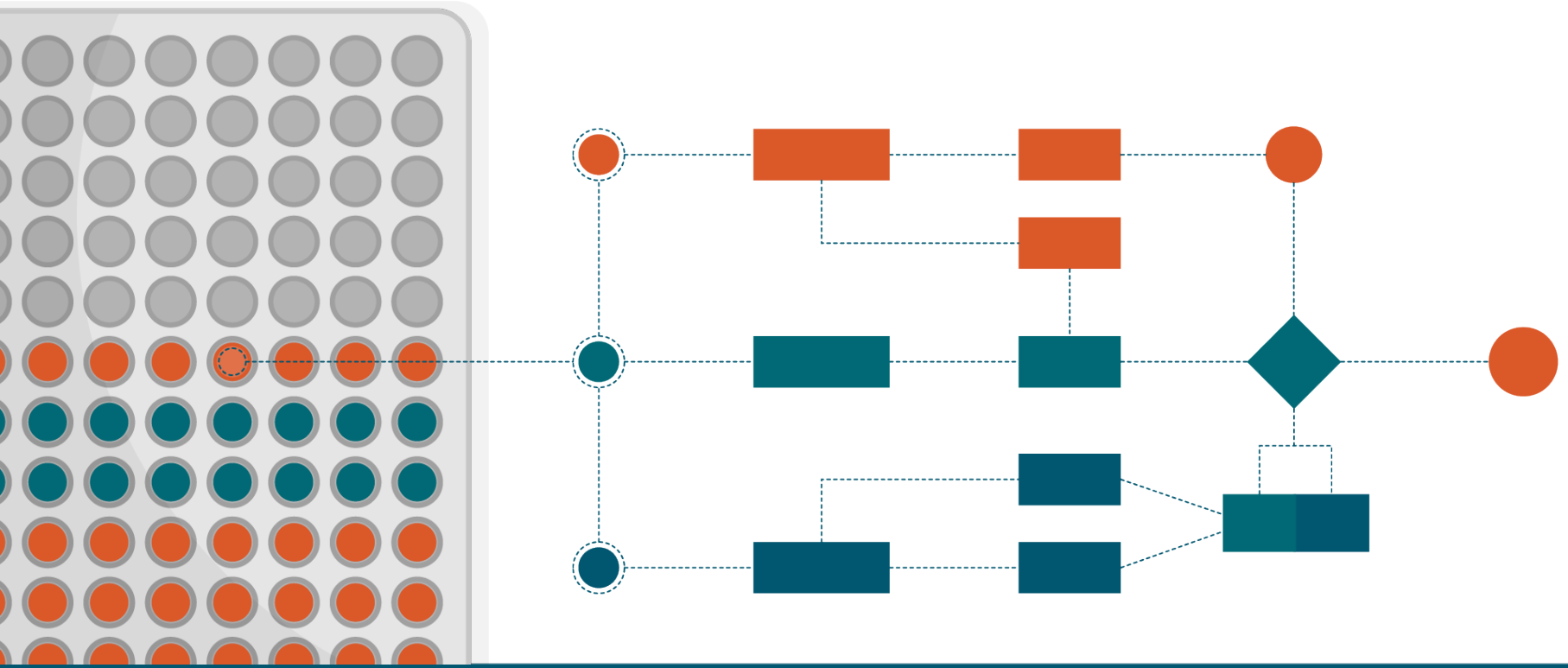
* Numbers in italics were not included in calculation for %CV or in graph below.

- **Sample 1** Not Useful as Numbers **BLQ**
- **Matrix Effect Observed at Neat**
- **Below Limit** of Quantification



Parallelism observed at a narrow window of **1:2 to 1:4 dilution**

Squeeze More From Your Sample with **Multiplexing**



Case Study 3: 45-plex

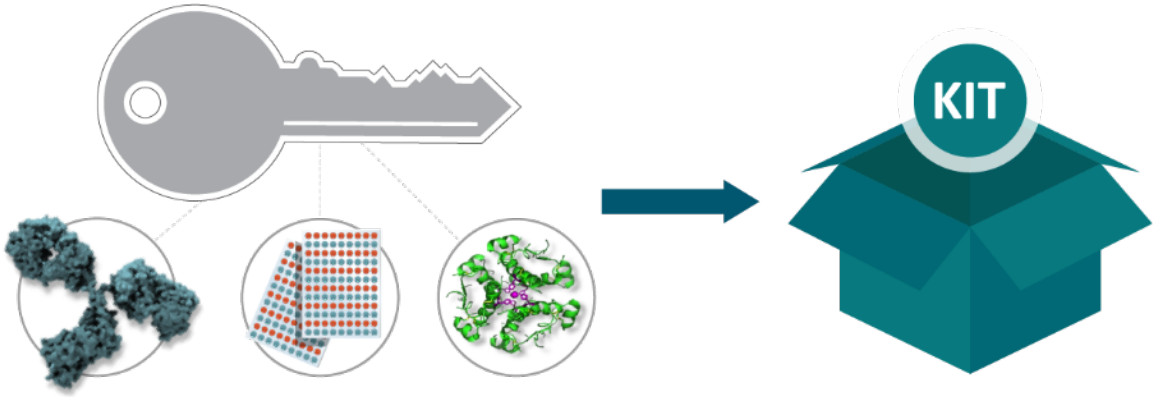
GM-CSF (44)	not linear
IFN-gamma (43)	not linear
TNF-ALPHA (45)	1 : 8
IL-10 (28)	not linear
IL-12p70 (34)	not linear
IL-13 (35)	1 : 8
IL-17A (36)	not linear
IL-18 (66)	not linear
IL-1beta (18)	not linear
IL-2 (19)	1 : 4
IL-21 (72)	not linear
IL-22 (76)	1 : 8
IL-23 (63)	not linear
IL-27 (14)	1 : 4
IL-4 (20)	not linear
IL-5 (21)	1 : 8
IL-6 (25)	1 : 8
IL-9 (52)	1 : 8
Eotaxin (33)	not linear
GRO-alpha (61)	not linear
IFN-alpha (48)	not linear

IL-15 (65)	1 : 8
IL-1alpha (62)	1 : 4
IL-1RA (38)	1 : 8
IL-31 (37)	not linear
IL-7 (26)	1 : 4
IL-8 (27)	not linear
IP-10 (22)	not linear
MCP-1 (51)	1 : 32
MIP-1alpha (12)	not linear
MIP-1beta (47)	1 : 4
RANTES (42)	not linear
SDF-1alpha (13)	not linear
TNF-beta (54)	not linear
BDNF (57)	not linear
LIF (15)	not linear
SCF (39)	1 : 4
VEGF-D (53)	1 : 8
bNGF (55)	not linear
EGF (56)	not linear
FGF-2 (75)	not linear
HGF (46)	1 : 8
PDGF-BB (77)	1 : 8
PlGF-1 (29)	1 : 4
VEGF-A (78)	1 : 32

- ✓ **25 out of 45 biomarkers** did not show any parallelism.
- ✓ **7 biomarkers** required at least 1:4 dilutions, 11 biomarkers required at least 1:8 dilutions and 2 required at least 1:32 dilutions before start seeing acceptable parallelism.
- ✓ **Challenge:** if try to analyze all samples at once at 1:32 dilution, then the assay is not sensitive for most of the biomarkers; if run as 1:4 or 1:8 dilutions, then require running the samples at least 3 separate dilutions.

Key to Setting Up a **“Solid”** Single and Multiplex LBA

Critical reagents are those essential components of LBAs whose unique characteristics are crucial to assay performance and therefore require thorough characterization and documentation.



Ligand Binding Assays in the 21st Century Laboratory: Recommendations for Characterization and Supply of Critical Reagents.

Denise M. O’Hara, Valerie Theobald, Adrienne Clements Egan, Joel Usansky, Murli Krishna, Julie TerWee, Mauricio Maia, Frank P. Spriggs, John Kenney, Afshin Safavi, and Jeannine Keefe

AAPS Journal, Volume 14, Number 2: (2012), 316-328

Case Study

1

3 SinglePlex Assay Kits and 1 custom 8-plex Assay Kit were validated for quantification of 11 analytes in human plasma ~1 year in advance of sample analysis.

2

Lot bridging studies were performed using stability samples and freshly prepared QCs run on old and new lots of kits (4 lots manufactured and tested over 3 years).

3

Ratios of the samples' mean concentrations between old and new lots of kits were examined to determine if a correction factor was needed to bridge measurements from different kit lots.

$$\frac{\text{Old}}{\text{New}}\text{Lot Ratio} = \frac{\text{Sample Mean concentration from Old Lot}}{\text{Sample Mean concentration fro New Lot}}$$

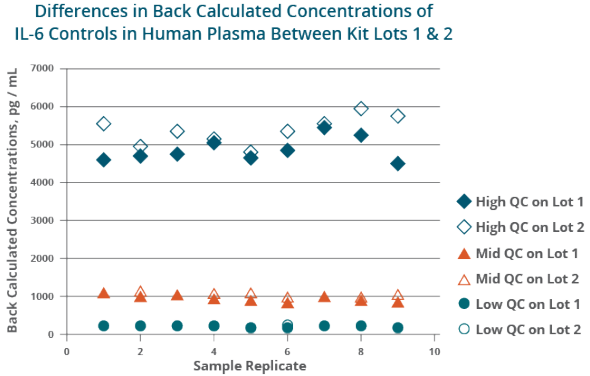
Case Study: Biomarker Analysis Requiring Correction Factors

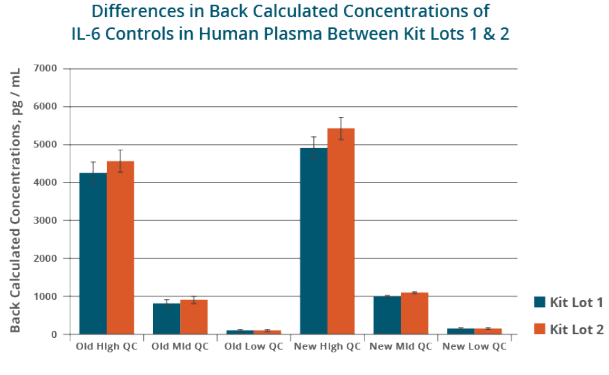
Kits	Analyte	Correction Factors		
		Lot 1 to Lot 2	Lot 2 to Lot 3	Lot 3 to Lot 4
Custom 8-Plex	TNF-alpha	N/A	N/A	N/A
	IFN-alpha	N/A	N/A	N/A
	IL-6	N/A	1.31	N/A
	IP-10	1.54	N/A	N/A
	I-TAC	1.29	N/A	N/A
	MCP-1	N/A	N/A	N/A
	TARC	N/A	N/A	N/A
	MIG	1.39	N/A	N/A
SinglePlex	MCP-2	N/A	N/A	2.02
	E-Selectin	N/A	N/A	N/A
	VEGF	N/A	0.65	N/A

Case Study (Analytes Not Requiring Correction Factors)

Kits	Analyte	Correction Factors		
		Lot 1 to Lot 2	Lot 2 to Lot 3	Lot 3 to Lot 4
Custom 8-Plex	TNF-alpha	N/A	N/A	N/A
	IFN-alpha	N/A	N/A	N/A
	IL-6	N/A	1.31	N/A
	IP-10	1.54	N/A	N/A
	I-TAC	1.29	N/A	N/A
	MCP-1	N/A	N/A	N/A
	TARC	N/A	N/A	N/A
	MIG	1.39	N/A	N/A
SinglePlex	MCP-2	N/A	N/A	2.02
	E-Selectin	N/A	N/A	N/A
	VEGF	N/A	0.65	N/A

IL-6 From First Lot Bridging of Custom 8-Plex Kit in Which **No Correction Factor is Required**



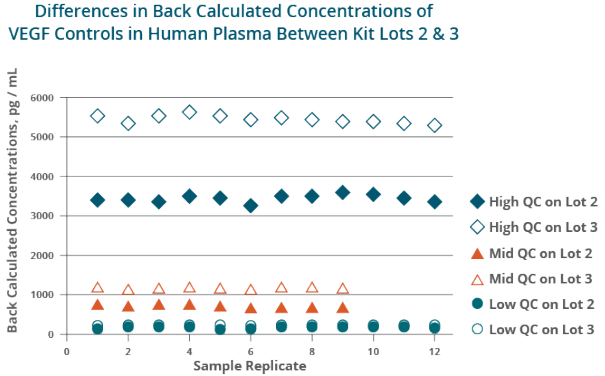


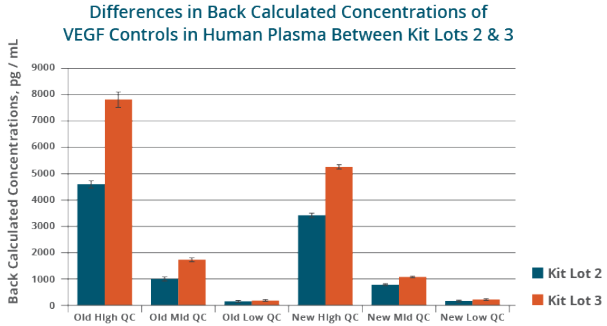
Sample	IL-6 Kit Lot 1 (n=9)			IL-6 Kit Lot 2 (n=9)			Old/New Kit Ratio
	Mean	Std Dev	CV (%)	Mean	Std Dev	CV (%)	
Old High Stability Sample	4258	252	5.9	4607	328	7.1	0.924
Old Mid Stability Sample	852	56.8	6.7	930	49.7	5.3	0.916
Old Low Stability Sample	167	12.5	7.5	174	17.8	10.2	0.960
New High QC	4915	297	6.0	5401	377	7.0	0.910
New Mid QC	961	71.7	7.5	1047	38.9	3.7	0.918
New Low QC	188	13.1	7.0	197	7.9	4.0	0.954

Case Study (Analytes Requiring Correction Factors)

Kits	Analyte	Correction Factors		
		Lot 1 to Lot 2	Lot 2 to Lot 3	Lot 3 to Lot 4
Custom 8-Plex	TNF-alpha	N/A	N/A	N/A
	IFN-alpha	N/A	N/A	N/A
	IL-6	N/A	1.31	N/A
	IP-10	1.54	N/A	N/A
	I-TAC	1.29	N/A	N/A
	MCP-1	N/A	N/A	N/A
	TARC	N/A	N/A	N/A
	MIG	1.39	N/A	N/A
SinglePlex	MCP-2	N/A	N/A	2.02
	E-Selectin	N/A	N/A	N/A
	VEGF	N/A	0.65	N/A

VEGF From Second Lot Bridging of Custom 8-Plex Kit in Which a Correction Factor is Required

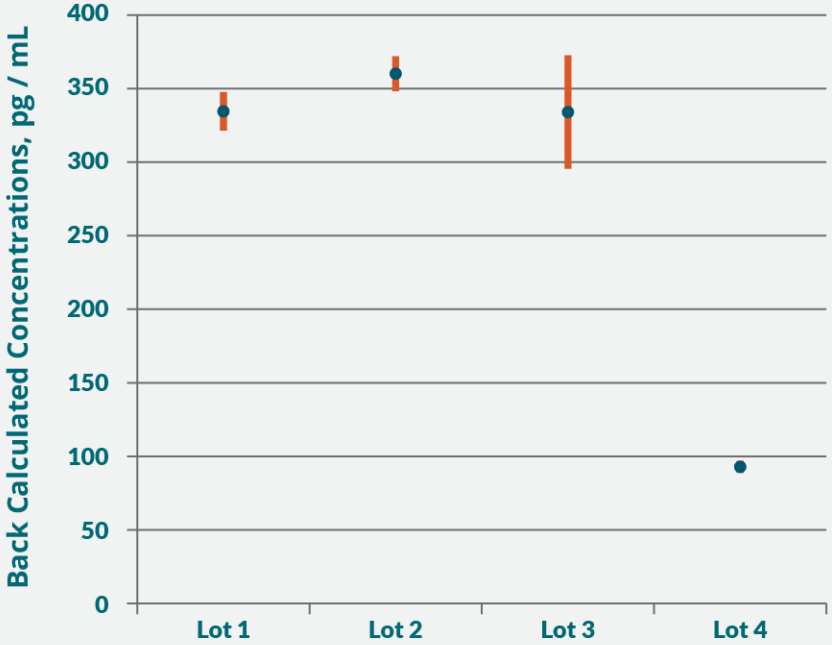




Sample	Old VEGF Kit Lot (n=9-12)			New VEGF Kit Lot (n=9-12)			Old/New Kit Ratio	Correction Factor
	Mean	Std Dev	CV (%)	Mean	Std Dev	CV (%)		
Old High QC	4611	106	2.3	7704	361	4.7	0.599	0.650
Old Mid QC	1008	22.4	2.2	1627	58.2	3.6	0.620	
Old Low QC	186	7.26	3.9	270	10.8	4.0	0.689	
New High QC	3438	112	3.3	5398	91.4	1.7	0.637	
New Mid QC	756	30.0	4.0	1165	22.5	1.9	0.649	
New Low QC	137	17.2	12.5	194	5.51	2.8	0.707	

If Kit Lot Bridging **Had Not** Been Performed **for IP-10 (8-Plex)**

If Kit Lot Bridging **Had Not** Been Performed **for MCP-2 (SinglePlex)**



Challenges of Multiplex Kit Lot Bridging

- Challenge topic
- May result in initial project cost increase
- Some may resist the idea
- It is needed
- One solution may not be for all

Conclusions

1. For biomarker assays, the calibrators are typically either recombinant or purified materials and therefore most often not identical to the endogenous form being measured. Therefore, **the results for most RUO kits are based on "relative accuracy".**

2. Parallelism assessment is one of the key parameters for evaluating biomarker studies and **should be initiated early during the assay development stage.**

3. **For parallelism assessment, one solution does not fit all.** Need to put your science hat on, be systematic in your approach (not picking and choosing) and make sure the rationale behind your parallelism decision is well documented.

4. Proper design of lot bridging experiments that measure the effect of using different lots of immunoassay kits are **critical to ensuring there is consistent measurement of the analyte over the course of the study.**

Working Together in **True Partnership**