The Blind Leading the Clear-Sighted
Undermining the Science of Context of Use

EBF Biomarker Focus Workshop
27-28 April 2021

Lauren Stevenson

Overview

- Foundational Concepts
- Case Study – A devolution from Science to Box-checking
  - Risks and consequences
- Conclusions
- Perspective

Foundational Concepts: Validation

What is Validation?

- A process to establish that the performance of a test, tool or instrument is acceptable for its intended purpose (BEST)
- Method/Assay Validation
  - Method validation is the process used to confirm that the analytical procedure employed for a specific test is suitable for its intended use (Ludwig Huber, *Validation and Qualification in Analytical Laboratories)*
  - Assay validation provides an assurance of reliability during normal use and is sometimes referred to as "the process of providing documented evidence that the method does what it is intended to do" (www.fws.gov)
- **Validated = Fit for Purpose!**

What is Context of Use?

- BEST resource 2016:
  - The Context of Use (COU) is “A statement that fully and clearly describes the way the medical product development tool is to be used and the medical product development related purpose of the use”

- Or, simply…
  - Context of Use = The Purpose in Fit-For-Purpose

Validation Requires COU

- If Validated = Fit-for-Purpose, and
- And COU = the Purpose in FFP
- Then Validation = Fit-for-COU

- COU is requisite for validation

- **No COU, no validation!**

Foundational Concepts: First Principles Thinking

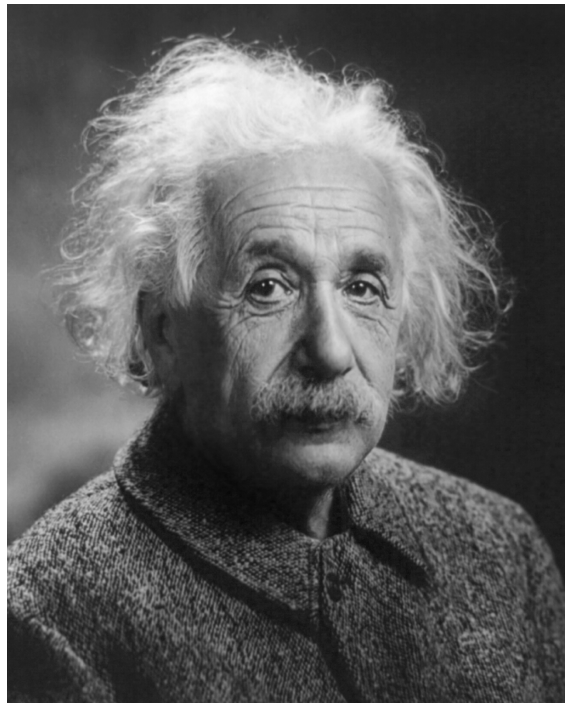
First Principles Thinking vs Reasoning by Analogy

- First Principles Thinking
  - Actively questioning everything you think you know about a given problem and then creating new knowledge and solutions from the ground up
  - #BeAScientist
- Reasoning by Analogy
  - Building knowledge and solving problems based on prior assumptions and beliefs, and perceived ‘best practices’
- Reasoning by Analogy tends to lead to bad decisions
  - Misapplication/overapplication *because it hasn’t been thought through*
  - Example: Misapplying PK Assay BMV Guidance to biomarker assays

Ref: https://medium.com/the-mission/elon-musks-3-step-first-principles-thinking-how-to-think-and-solve-difficult-problems-like-a-ba1e73a9f6c0

First Principles

"If I had an hour to solve a problem, I'd spend 55 minutes thinking about the problem and 5 minutes thinking about solutions" – Albert Einstein



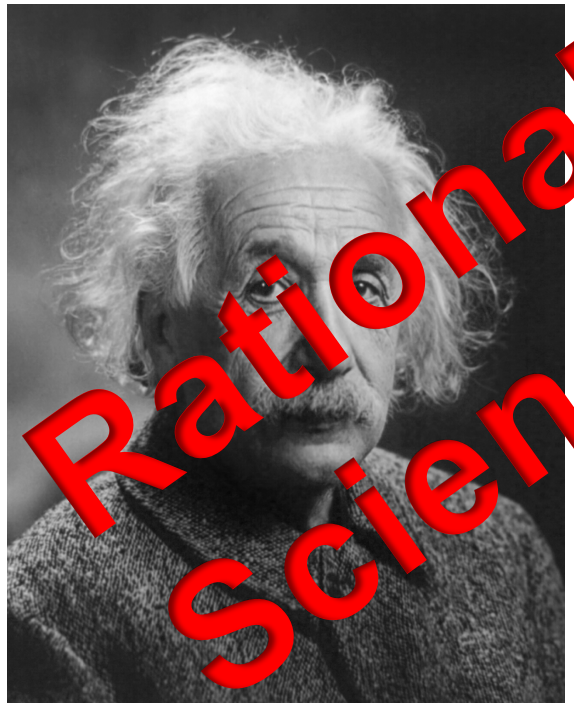
Reasoning by Analogy

"The person who says he knows what he thinks but cannot express it usually does not know what he thinks" – Mortimer Adler

- That's how we've always done it
- It's in the (BMV) guidance
- Because regulators might ask about it

First Principles

"If I had an hour to solve a problem, I'd spend 55 minutes thinking about the problem and 5 minutes thinking about solutions" – Albert Einstein



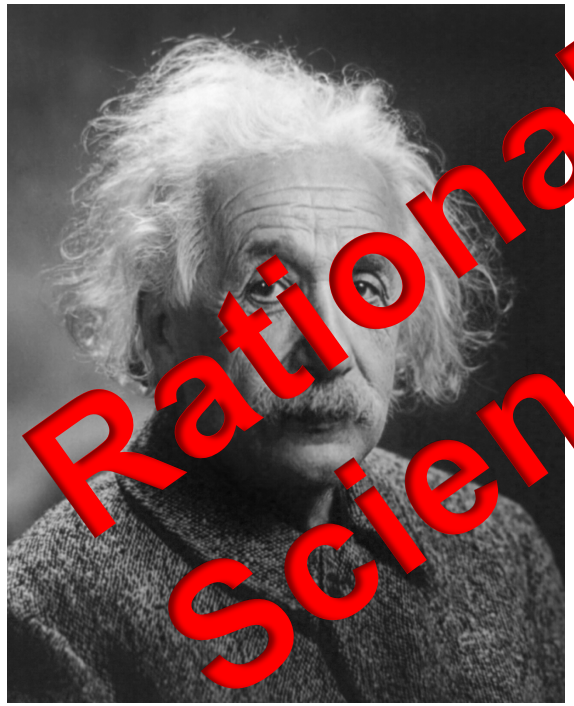
Reasoning by Analogy

"The person who says he knows what he thinks but cannot express it usually does not know what he thinks" – Mortimer Adler

- That's how we've always done it
- It's in the (BMV) guidance
- Because regulators might ask about it

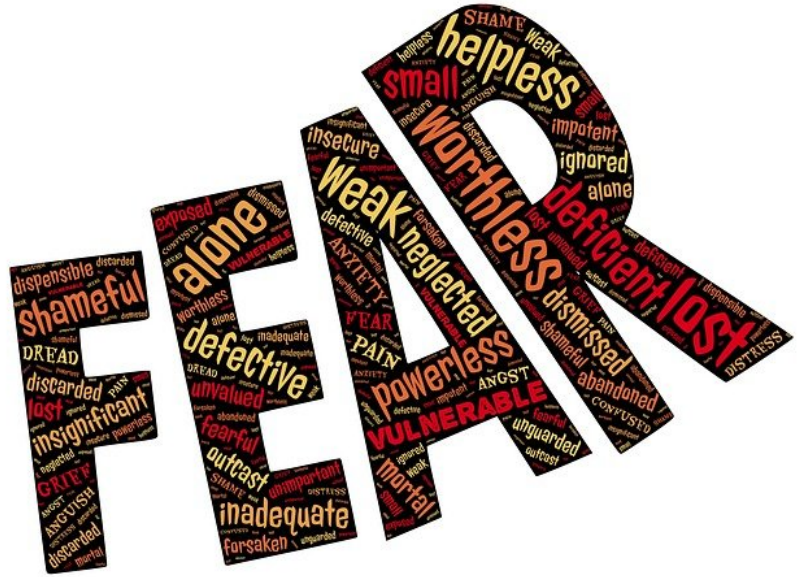
First Principles

"If I had an hour to solve a problem, I'd spend 55 minutes thinking about the problem and 5 minutes thinking about solutions" – Albert Einstein



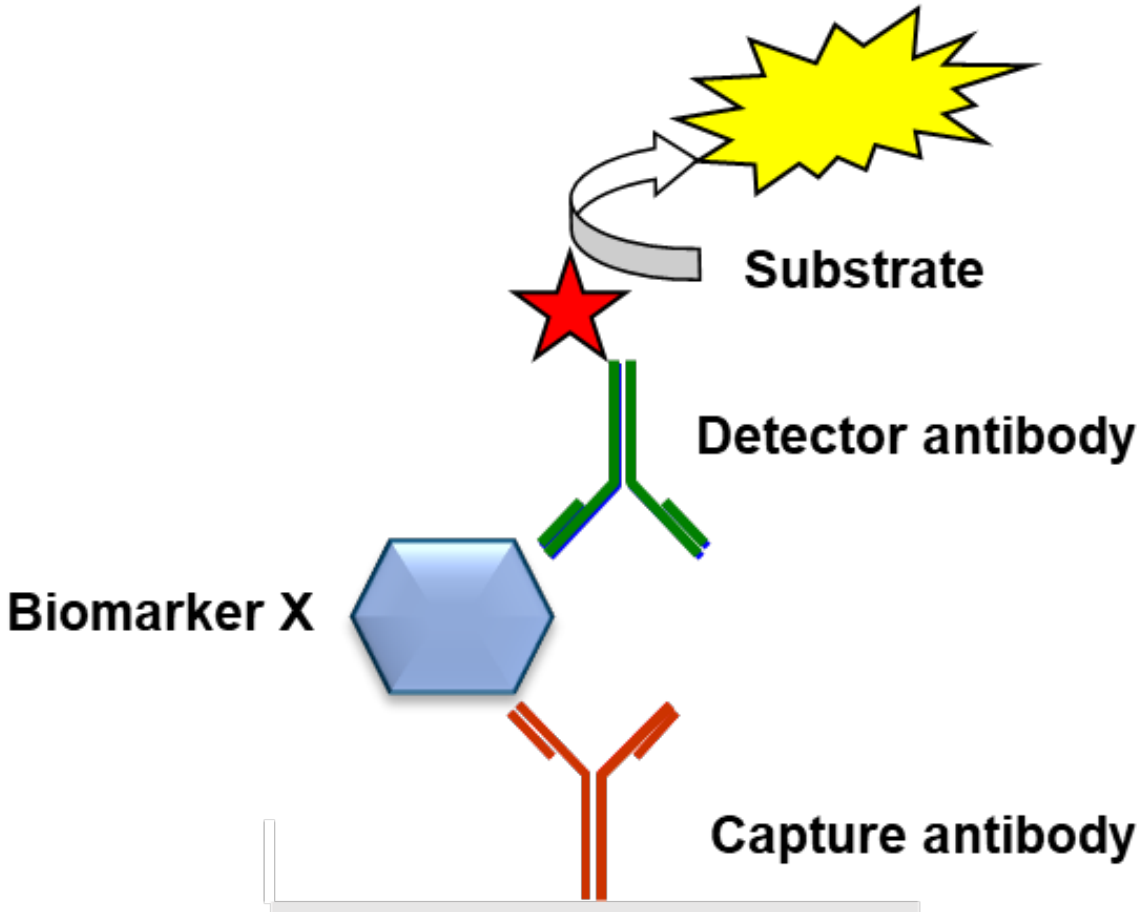
Reasoning by Analogy

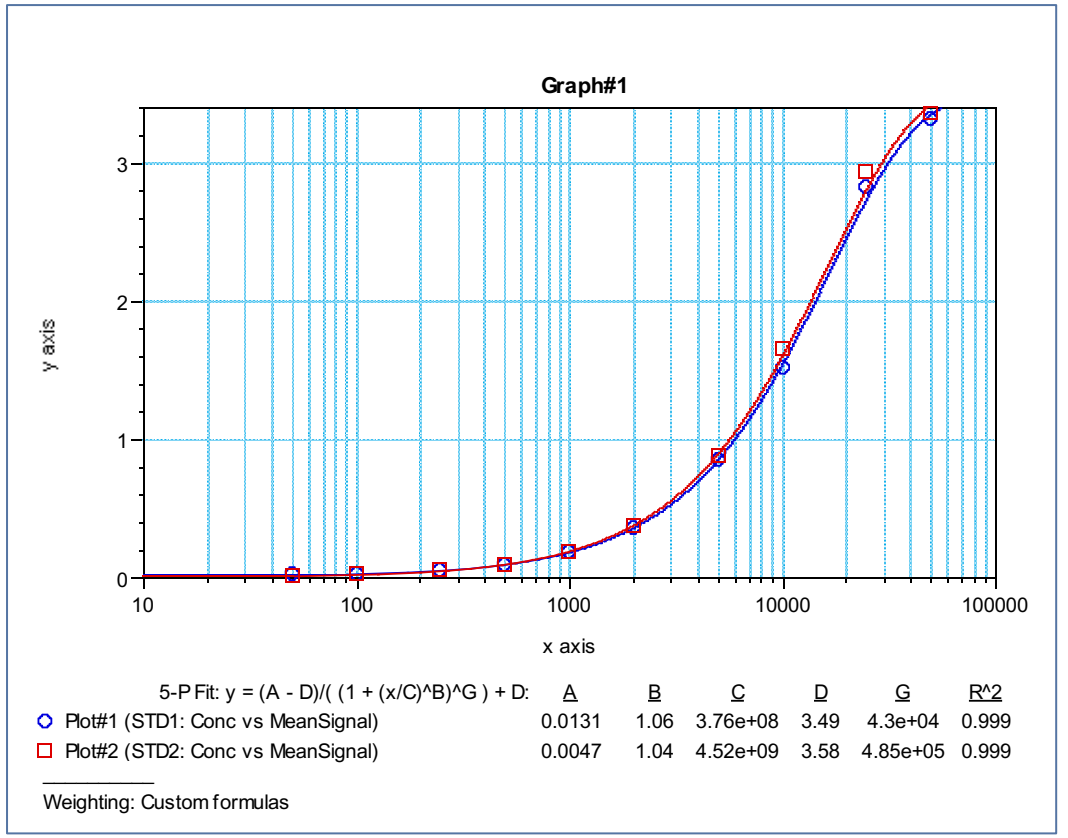
"The person who says he knows what he thinks but cannot express it usually does not know what he thinks" – Mortimer Adler



Case Study

Case Study: Sandwich ELISA for Biomarker X





Case Study COU

- If I told you, I'd have to kill you….

- "A specific COU for this assay has not yet been defined. The purpose of this validation is therefore to characterize the assay's analytical performance for relative quantification of Biomarker X in plasma and serum samples."
  - Initial studies would include Phase 0 studies to inform biological variability

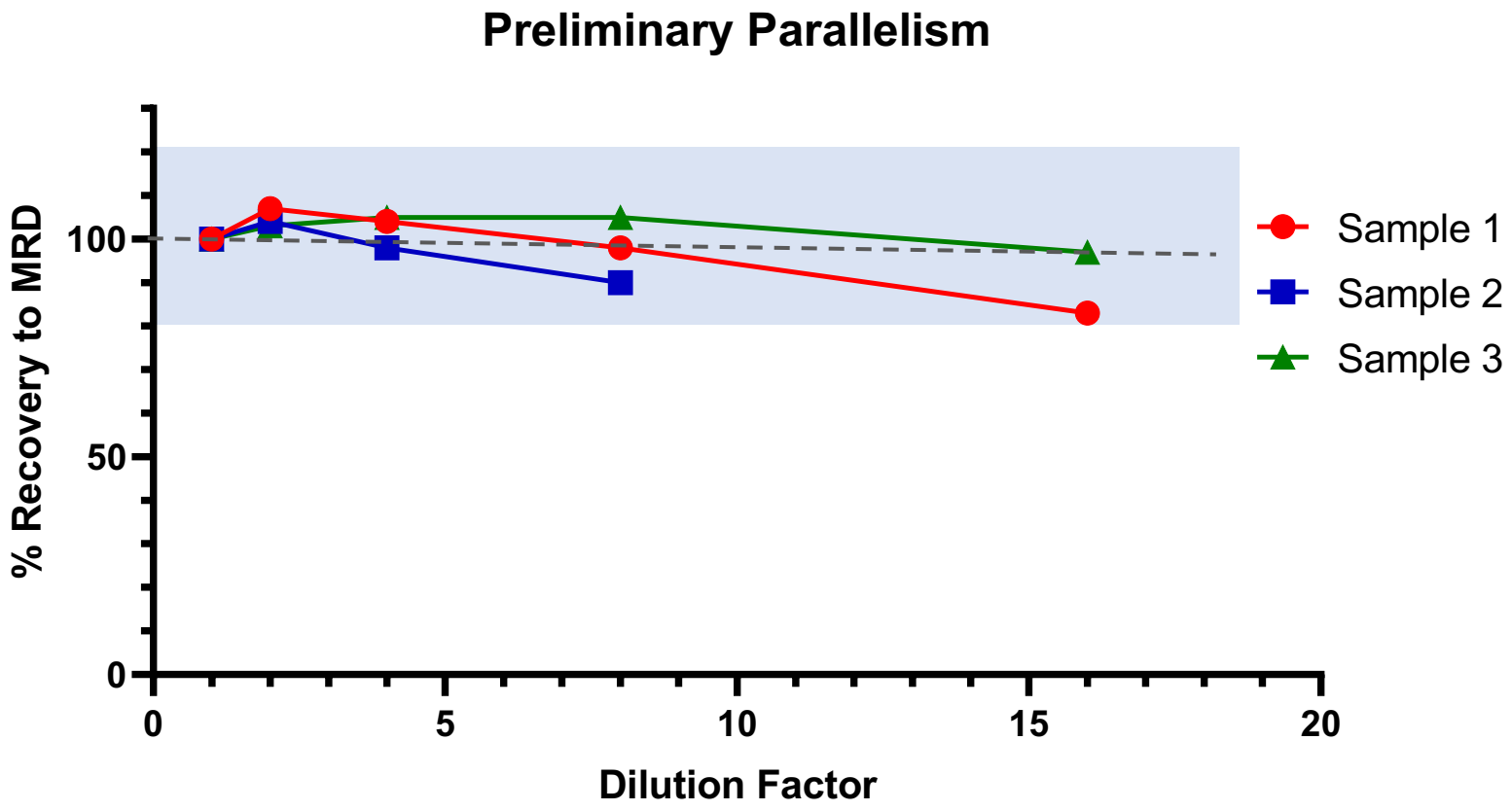
Case Study – Method Development Drivers

- Specificity – potential interfering analyte
- Sensitivity – 100 pg/mL target sensitivity
- Precision – better precision more likely to support potential COUs
- Performance across multiple sample matrices - to support sample analysis for a variety of banked studies

- Initial optimization experiments and subsequently, the entirety of assay development performed in singlet
- Why?
- Because……First Principles

Method Development: Preliminary Parallelism

- Samples are representative of study population
- Range of Biomarker X levels in presence of potential interfering analyte



Method Development Driver: Sensitivity

		MRD	MRD x2	MRD x4	MRD x8	MRD x16
Sample 1 N = 2	**Mean**	278	148	70	33	NA
	Stdev	2.1	2.1	3.5	1.4	NA
	CV	0.8	1.4	5.1	4.3	NA
Sample 2 N = 2	**Mean**	99	51	21	NA	NA
	Stdev	1.4	4.2	2.1	NA	NA
	CV	1.4	8.3	10.3	NA	NA
Sample 3 N = 2	**Mean**	802	408	210	104	48
	Stdev	19.8	4.9	3.5	0.7	0.7
	CV	2.5	1.2	1.7	0.7	1.5

- Estimated endogenous analyte LLOQ ~50 pg/mL
- Bonus: Preliminary look at intra-assay precision across multiple dilutions

Case Study – Method Development Drivers

- Sensitivity 
- Precision
- Multiple sample matrices

Buffer QC Inter-assay Precision (singlet)

Analysis Day	ULOQ 25,000 pg/mL	HQC 20,000 pg/mL	MQC 2500 pg/mL	LQC 75 pg/mL
Day 1	24852	19878	2240	86
Day 2	26075	20644	2361	80
Day 3	24837	19977	2416	85
Day 4	27699	23242	2617	85
Day 5	26396	22907	2647	84
Day 6	26478	21630	2374	101
Day 7	28897	23063	2761	84
Mean	26462.0	21620.1	2488.0	86.4
% RE	5.8	8.1	-0.5	15.2
Stdev	1463.2	1475.1	188.1	6.7
%CV	5.5	6.8	7.6	7.8
Total Error	11.0	14.9	8.1	23.0

Two analysts over 7 days, including qualification runs for Analyst 2

Endogenous QC Inter-assay Precision (singlet)

Analysis Day	High EQC (pg/mL)	*Low EQC (pg/mL)
Day 1	193	92
Day 2	193	97
Day 3	188	91
Day 4	193	110
Day 5	217	107
Mean	196.8	99.4
Stdev	11.5	8.7
%CV	5.8	8.7

*1/2 Dilution of High EQC, stored frozen

Case Study – Method Development Drivers

- Sensitivity 
- Precision 
- Multiple sample matrices

Matrix Comparison

Comparison	% Difference
Matrix 1 and 2	1.2
Matrix 1 and 3	5.6
Matrix 2 and 3	4.4

Sample	Analysis Day	Measured conc (pg/mL)	mean	Stdev	CV
Matrix 1	Day 1-1	233	242.5	10.6	4.4
	Day 1-2	234			
	Day 2-1	254			
	Day 2-2	249			
Matrix 2	Day 1-1	250	245.3	8.1	3.3
	Day 1-2	252			
	Day 2-1	245			
	Day 2-2	234			
Matrix 3	Day 1-1	256	256.0	16.7	6.5
	Day 1-2	272			
	Day 2-1	263			
	Day 2-2	233			
Overall mean		247.9			
Stdev		12.7			
CV		5.1			

Case Study – Method Development Drivers

- Sensitivity 
- Precision 
- Multiple sample matrices 

Conclusions from Assay Development

- Sensitivity
  - Parallelism demonstrates target LLOQ of 100 pg/mL achievable, with 50 pg/mL possible
- Precision
  - Assay demonstrates high performance in singlet: Inter-assay CVs ≤ 8.7%
- Different sample types can be evaluated in the same assay
  - % difference between 3 different matrices ≤ 5.6%
- Other (data not shown)
  - Assay tolerates high levels of potential interfering analyte
  - Parallelism demonstrated across multiple individuals and matrices = Selectivity
  - Preliminary stability for EQC
  - Robustness and ruggedness – multiple preps, multiple analysts, multiple reagent lots, etc.

Validation Plan v1 – Science-driven

- Calibration Curve performance
- Intra- and inter-assay precision and relative accuracy
  - 4 levels of buffer QCs to cover curve range; 3 levels of EQCs to cover anticipated sample concentration range
- Parallelism
  - Confirm MRD, Selectivity, Estimation of endogenous analyte LLOQ
- Specificity - Confirm non-interference
- Sample matrix – Confirm comparability of 3 matrices
- Stability
- Robustness and Ruggedness

Validation Plan v1 – Acceptance criteria

- Science-driven – based on assay performance data
- Statistically determined to define expected performance ranges
- Can then be used to define what can (and can't) be concluded from assay data and guide utility for future potential COUs

Validation Plan v2: Modified by client management

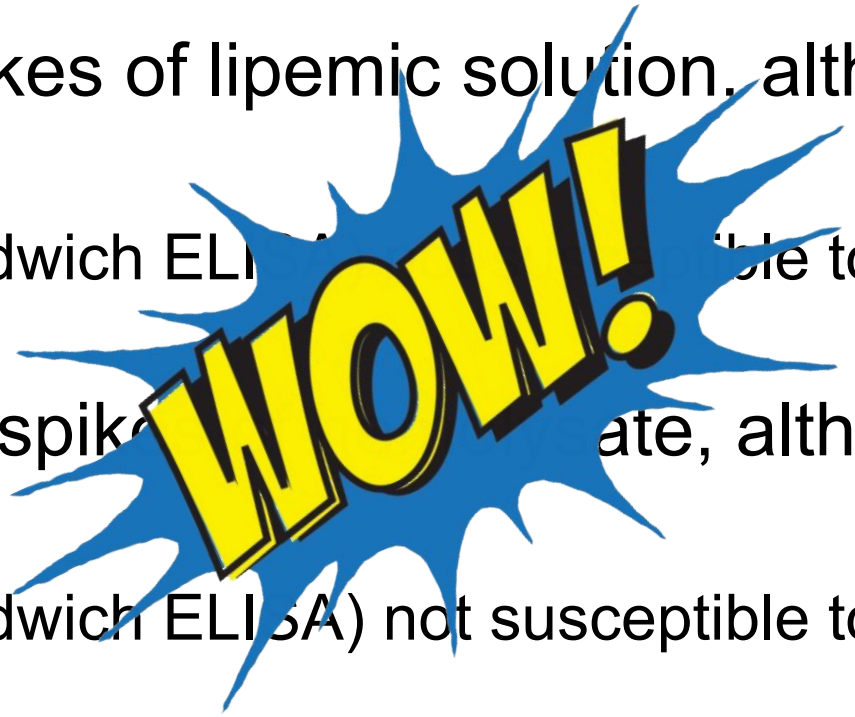
Everything in Plan v1, PLUS:

- Duplicate analysis
- Lipemic Samples – 3 replicate spikes of lipemic solution, although…
  - Disease population not lipemic
  - Assay methodology (sequential sandwich ELISA) not susceptible to interference from lipemia
- Hemolyzed samples – 3 replicate spikes of hemolysate, although…
  - Hemolyzed samples rarely occur
  - Assay methodology (sequential sandwich ELISA) not susceptible to interference from hemolysis
- Pre-set 25% acceptance criteria

Validation Plan v2: Modified by client management

Everything in Plan v1, PLUS:

- Duplicate analysis
- Lipemic Samples – 3 replicate spikes of lipemic solution, although…
  - Disease population not lipemic
  - Assay methodology (sequential sandwich ELI[illegible]ible to interference from lipemia
- Hemolyzed samples – 3 replicate spik[illegible]ate, although…
  - Hemolyzed samples rarely occur
  - Assay methodology (sequential sandwich ELISA) not susceptible to interference from hemolysis
- Pre-set 25% acceptance criteria



Validation Plan v2 – Why?

- “Our organization is very risk averse”
- Organizational leaders are rarely biomarker development experts
  - Lack of direct experience engaging with Regulators on biomarkers
- Reasoning by analogy - belief that regulators require following BMV

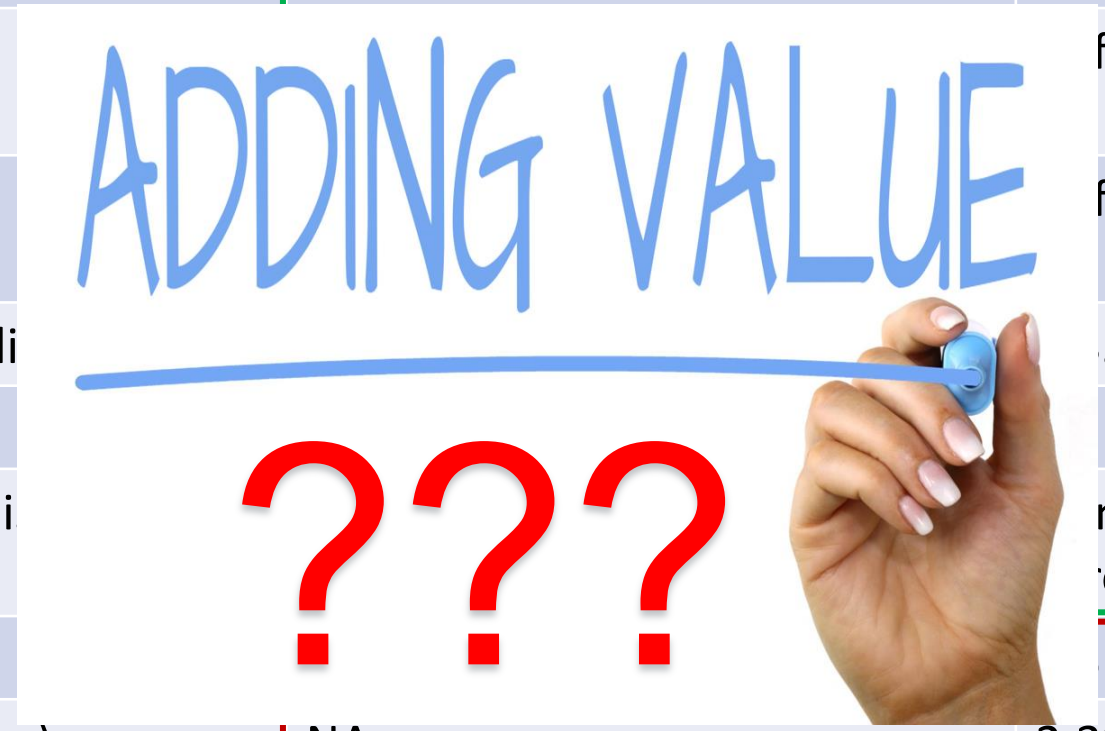
FEAR

Assay Performance: Method Development vs. Validation

Parameter	Method Development (Singlet)	Validation (Duplicate)
Sensitivity (estimated LLOQ via parallelism)	~ 50 pg/mL	54.2 pg/mL
Intra-assay precision	NA	Buffer QCs: ≤ 6.4% EQCs: ≤ 4.4%
Inter-assay precision	Buffer QCs: ≤ 7.8% EQCs: ≤ 8.7%	Buffer QCs: ≤ 6.0% EQCs: ≤ 9.6%
Multiple matrices (% difference)	≤ 5.6%	≤ 8.1%
Specificity	No interference	No interference
Selectivity (via parallelism)	Demonstrated parallelism across individuals	Demonstrated parallelism across individuals
Lipemia (% difference)	NA	1%
Hemolysis (% difference)	NA	2.3%

Assay Performance: Method Development vs. Validation

Parameter	Method Development (Singlet)	Validation (Duplicate)
Sensitivity (estimated LLOQ via parallelism)	~ 50 pg/mL	54.2 pg/mL
Intra-assay precision		ffer QCs: ≤ 6.4% EQCs: ≤ 4.4%
Inter-assay precision		ffer QCs: ≤ 6.0% EQCs: ≤ 9.6%
Multiple matrices (% di		.1%
Specificity		interference
Selectivity (via paralleli		monstrated parallelism oss individuals
Lipemia (% difference)		
Hemolysis (% difference)	NA	2.3%



Precision Acceptance Criteria

Control (pg/mL)	Inter-assay %CV	Acceptance Criteria	3 x CV criteria	Potential consequences
LQC (100)	6.0	≤ 25%	≤ 18%	Acceptance of runs where buffer QCs are outside performance expectations (>4 SD)
MQC (4000)	3.5	≤ 25%	≤ 10.5%	
HQC (20000)	5.6	≤ 25%	≤ 16.8	
ULOQ (25000)	6.0	≤ 25%	≤ 18%	
eLLOQ (54.2)*	11.0	≤ 25%	≤ 33%	Rejection of runs where EQCs are within performance expectations (3 SD). Unnecessary sample repeats. Potential loss of meaningful low concentration data
eLQC (105)*	9.6	≤ 25%	≤ 28.8%	
eHQC (217)*	9.6	≤ 25%	≤ 28.8%	

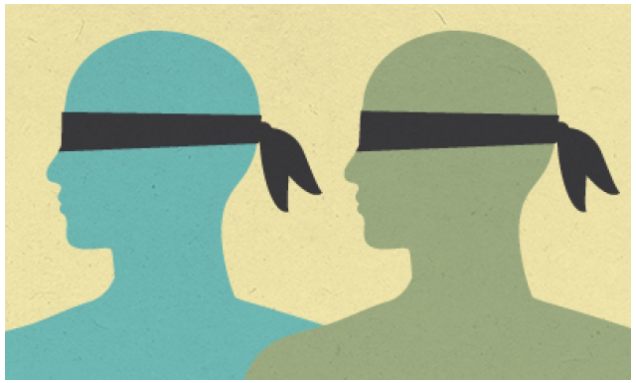
*EQC nominal concentrations set as mean of measured concentrations during validation

The Real Risks and Consequences

- Costs of duplicate analysis – resources, time and money
  - Double reagents/materials
  - Double bridging reagents lots
  - Half throughput, potentially double sample testing timelines
- Costs of non-scientific criteria
  - Misapplication of the assay and misinterpretation of data
    - Assumption that assay is only good for ‘exploratory’ work - due to application of arbitrary 25% ‘exploratory’ criteria?
    - Or…assumption that the assay is good for any COU – because it included all the ‘PK’ analyses?
- Perpetuation of non-science, preventing scientists from being scientists = Undermining Science
- Not simply a lack of adding value, but risk of getting it wrong

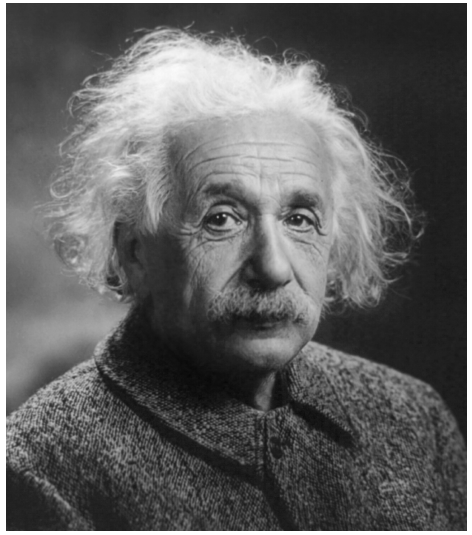
Conclusions

- This case study demonstrates how the stakeholder's stakeholders (leadership) can derail COU-driven biomarker assay validation



- In order to make lasting progress, we will need to find a means to educate and influence the layers of stakeholders within drug development organizations

- Where Regulators lead, industry will follow...



"The definition of insanity is doing the same thing over and over again but expecting different results."

Why do these non-scientific, non-value-added practices continue?

#BeAScientist

If there's no scientific rationale, it's not science

#BeAScientist

Critical thought partners: John Allinson, Erin Anderson