Evaluating Your Test Methods, Planning your QC: It's NOT a Regulatory Exercise!

Pat Garrett, Ph.D., DABCC

pgarrett@seracare.com

SeraCare Life Sciences, Inc.

AMP Annual Meeting

November 19, 2009

Orlando, Florida

It is, too, a regulatory exercise!

Okay, but it should be more, especially for molecular -- why?

- Relatively new technology, constantly evolving
- New analytes, new specimen types, appear daily – where are the standards and controls?
- Many manual steps, more room for operator modifications, errors
- Many tests are done once, and the patient has a diagnosis for life
- Many tests are multiplex (looking for multiple mutations)

SeraCare
LIFE SCIENCES

What are your expectations for your new test?

- Accurate
- Reproducible
- Robust
- Available
- Passes proficiency testing
- Anything else?

Call these Performance Characteristics – how do you evaluate them?

SeraCare
LIFE SCIENCES

However, a little guidance never hurts…

Begin by thinking about your expectations for the test, and about what could go wrong:

- Use your team's own knowledge and experience
- Read the package insert (aka IFU – Instructions for Use)
- Start a FMEA chart
  - What can happen?
    - How bad is it?
    - How likely is it?
- Make a plan to check for and mitigate the most likely and most severe problems
- Begin considering a QA/QC program to monitor for these

SeraCare
LIFE SCIENCES

What's a FMEA?

FMEA stands for 'failure mode effects analysis'

This is industrial quality-speak to describe **brainstorming ahead of time these questions:**

- What could go wrong with a process (aka test method)?
- How bad would that be?
- How likely is it to happen?
- Should you try to prevent this? If so, how?

The FMEA 'product' is a chart that helps to guide your further work in understanding and evaluating the test method.

SeraCare
LIFE SCIENCES

FMEA Setup

FMEA#						
Revision:						
Product/Process Name:	HIV PCR					
Prepared by:						
FMEA Team:						
Scoring	**Severity**		**Occurrence**	**Detection**	**Rating**	
	5 – Life Threatening Event		5 – Very High (>1/10)	5 – None (<1%)	**Critical**	>50
	4 – False negative, false positive, or similar performance related problems		4 – High (1/100)	4 – Low (1% -35%)	**Major**	18-50
	3 – Lab considers failure mode to be significant		3 – Moderate (1/1,000)	3 – Moderate (35% - 65%)	**Minor**	<18
	2 – Lab notices moderate problem		2 – Low (1/10,000)	2 – Likely (65% - 99%)		
	1 – Lab not aware		1 – Remote (<1/100,000)	1 – Certain (>99%)		

SeraCare
LIFE SCIENCES

Example FMEA (partial)

Design/Process Input	Potential Failure Mode	Potential Effects of Failure	Severity	Potential Cause of Failure	Occurrence	Controls	Detection	RPN = S x O x D	Risk Control Measures	References	Severity	Occurrence	Detection	RPN = S x O x D
Specimen Preparation														
Add sample to lysis buffer	Incomplete lysis	False negative	4	Operator or sample properties	2	None (kit controls are DNA-based not virus-based and do not test for lysis)	5	40	Add independent run control that is whole virus-based or demonstrate sample lysis is robust	Control Package Insert or sample validation	4	2	2	16
Incubate	Incomplete lysis	False negative	4	Operator or sample properties	2	None (kit controls are DNA-based not virus-based and do not test for lysis)	5	40	Add independent run control that is whole virus-based or demonstrate sample lysis is robust	Control Package Insert or sample validation	4	2	2	16
Centrifuge	Incomplete spin leaves contaminating cell debris which inhibits PCR	False negative	4	Operator or sample properties	2	None (kit controls are DNA-based not virus-based and do not test for lysis)	5	40	Add independent run control that is whole virus-based or demonstrate sample lysis is robust	Control Package Insert or sample validation	4	2	2	16

SeraCare
LIFE SCIENCES

Demonstrating that risk mitigation can work

Sabotage Condition	Replicate	CT Kit Controls	Prototype CT Control
	1	POS	NEG
	2	POS	NEG
	3	POS	NEG
	4	POS	NEG
No Lysis	5	POS	NEG
	1	POS	EQ
	2	POS	NEG
	3	POS	EQ
	4	POS	NEG
Omit 10 min incubation	5	FAIL	NEG
	1	POS	NEG
	2	POS	POS
	3	POS	POS
	4	POS	NEG
Incomplete neutralization	5	POS	EQ

SeraCare
LIFE SCIENCES

Verification vs. Validation vs. Establishment

Ongoing confusion as to precise definitions

- CLIA uses establishment and verification
- ISO has very similar definitions for validation and verification
- CLSI documents vary in their use of internationally established definitions

For purposes of our discussion:

- **Establishment:** design and development of performance characteristics
- **Validation**: determination of performance characteristics, once developed
- **Verification**: confirmation of performance characteristics previously determined by the manufacturer during validation
  - Verification studies smaller, narrower in scope, but extremely important!

This is how the test operates in your lab.

SeraCare LIFE SCIENCES

Verifying test performance in your lab

- **Accuracy**: can the test produce the correct result?
- **Reproducibility/precision**: can it do this consistently?
- **Robustness**: can all our techs run this method reliably on our schedule?
- **Reference range/reportable range**: was the manufacturer's range established with a population similar to ours?
- **Calibration plan**: what does the manufacturer recommend, and can we do this?
- **QC plan**: how do we detect errors when they occur?

SeraCare
LIFE SCIENCES

Guidance (beyond this workshop)

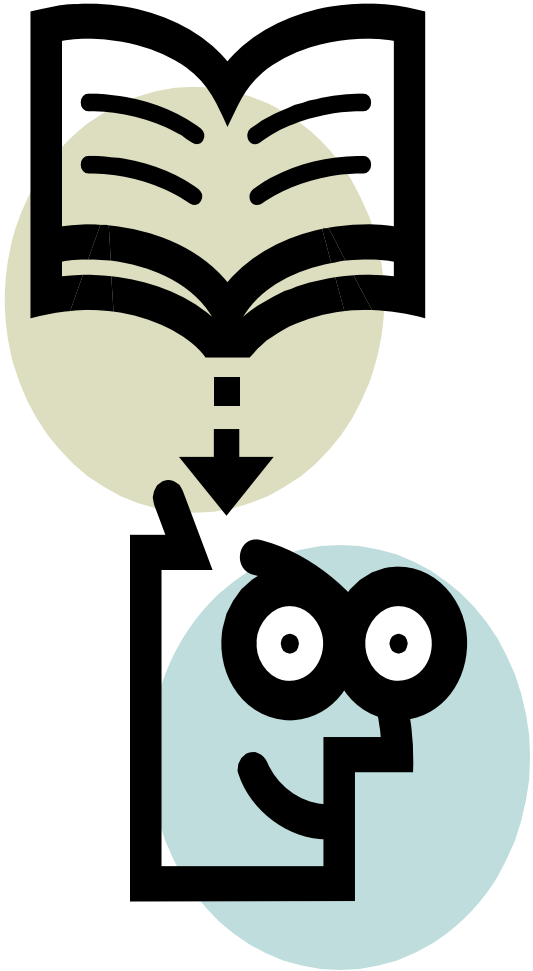
EP15-A2
User Verification of Performance for Precision and Trueness; Approved Guideline - Second Edition

EP12-A2
User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline - Second Edition

www.clsi.org

CLINICAL AND LABORATORY STANDARDS INSTITUTE ™

(Formerly NCCLS)
Providing NCCLS standards and guidelines, ISO/TC 212 standards, and ISO/TC 76 standards



SeraCare
LIFE SCIENCES

More guidance

Basic Method Validation, 3rd edition

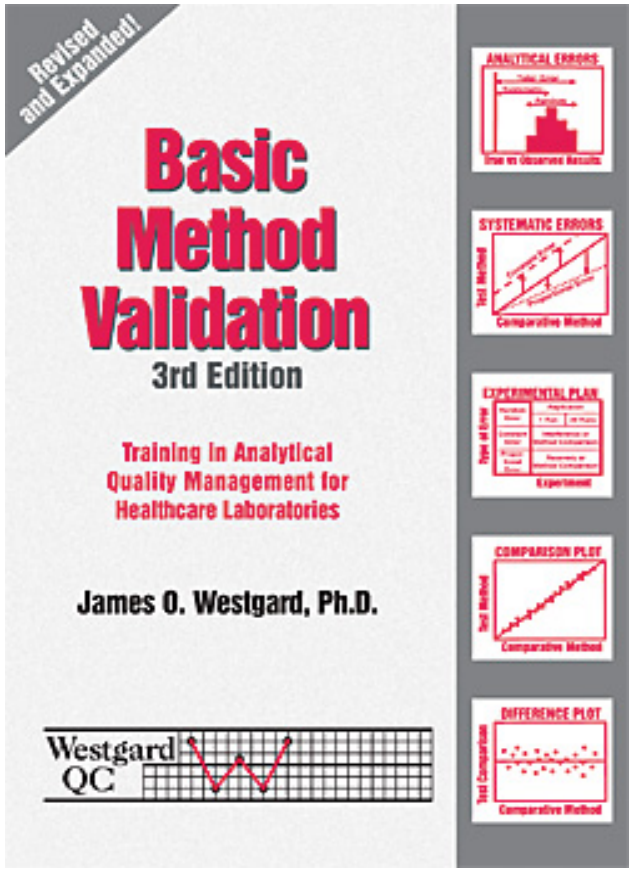
Training in Analytical Quality Management for Healthcare Laboratories

James O. Westgard, Ph.D.

www.westgard.com

EP Evaluator Release 9

Software for method evaluation from dgrhoads.com



SeraCare LIFE SCIENCES

Verification Studies Summary

Purpose: to assess error inherent in the test, so that you can recognize unacceptable changes

Studies:

- Accuracy
  - comparison of methods experiment
    - sensitivity, specificity, PPV, NPV
  - linearity
- Precision
  - within-run (repeatability)
  - between-run (reproducibility)
- Analytical sensitivity (detection limits)
- Analytical Specificity

SeraCare
LIFE SCIENCES

A quick review of error

There are only two kinds! (Well, three, actually.)

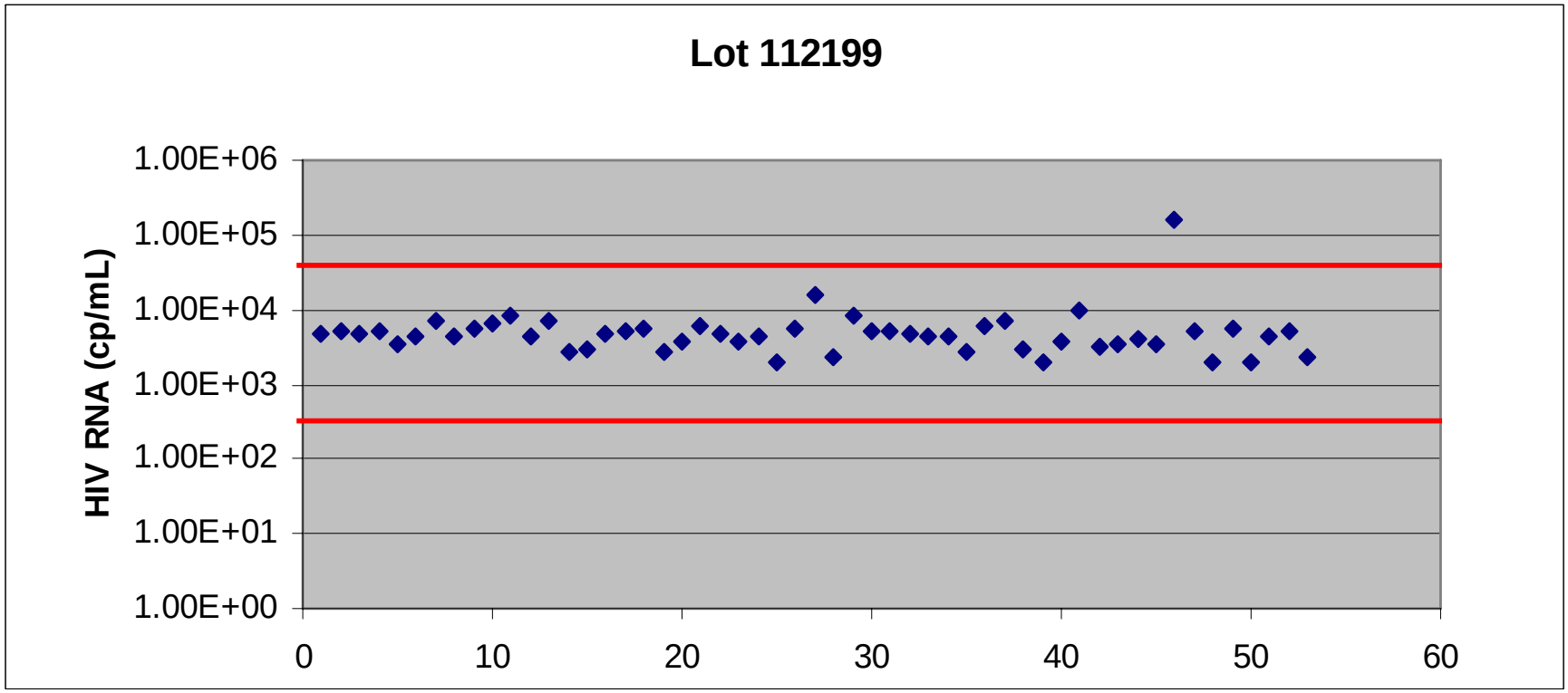
- **Random error**: an increase in the standard deviation (sd)
- **Systematic error**: a shift in the mean

The baseline settings for mean and sd are assessed in verification studies.

A good QC program monitors the test for changes in these.

SeraCare
LIFE SCIENCES

Watching for random errors with logs

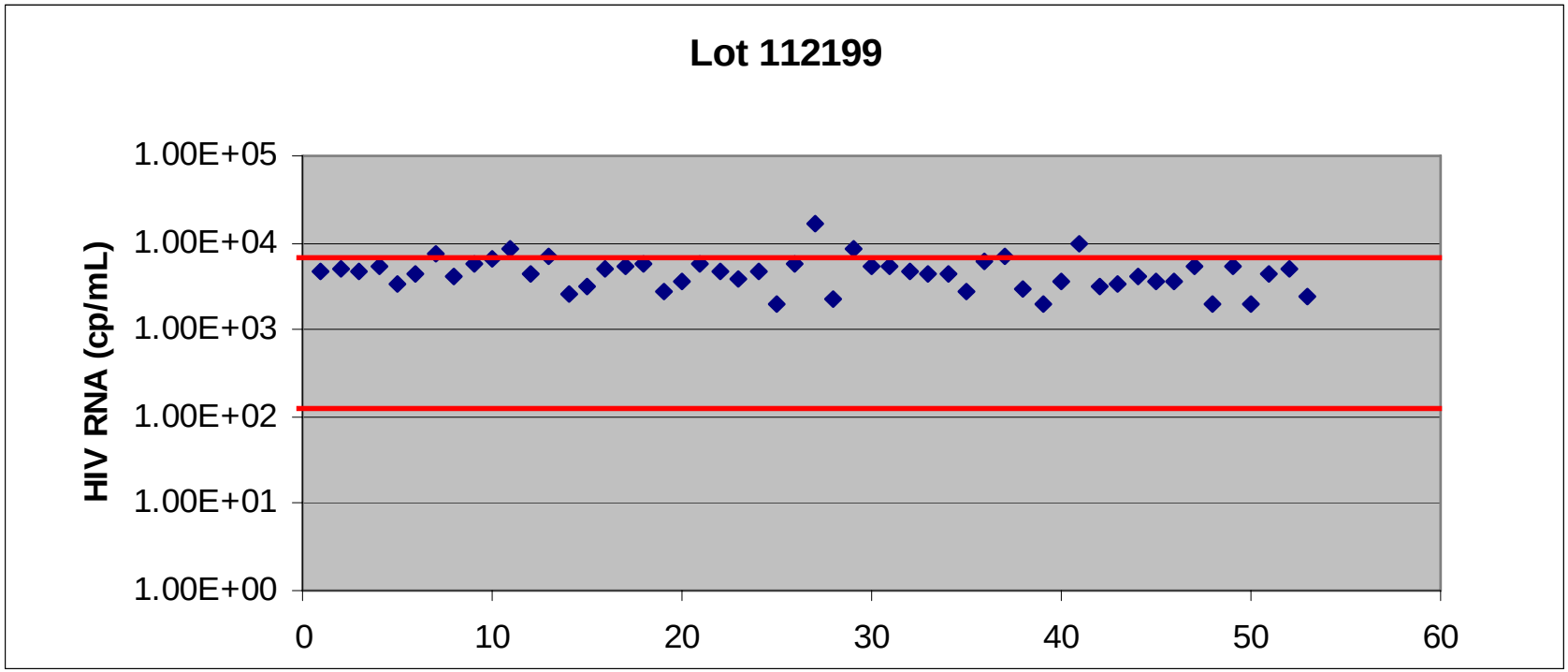


Mean = 7.65E+03, SD = 2.08E+04, CV = 271% (!)

Outlier is 155,000, values next to it are 3530 and 5210

SeraCare LIFE SCIENCES

When only the outlier is removed…



Mean is now 4.79E+03, SD = 2.36E+03, CV = 49%

SeraCare
LIFE SCIENCES

An error of the third kind…

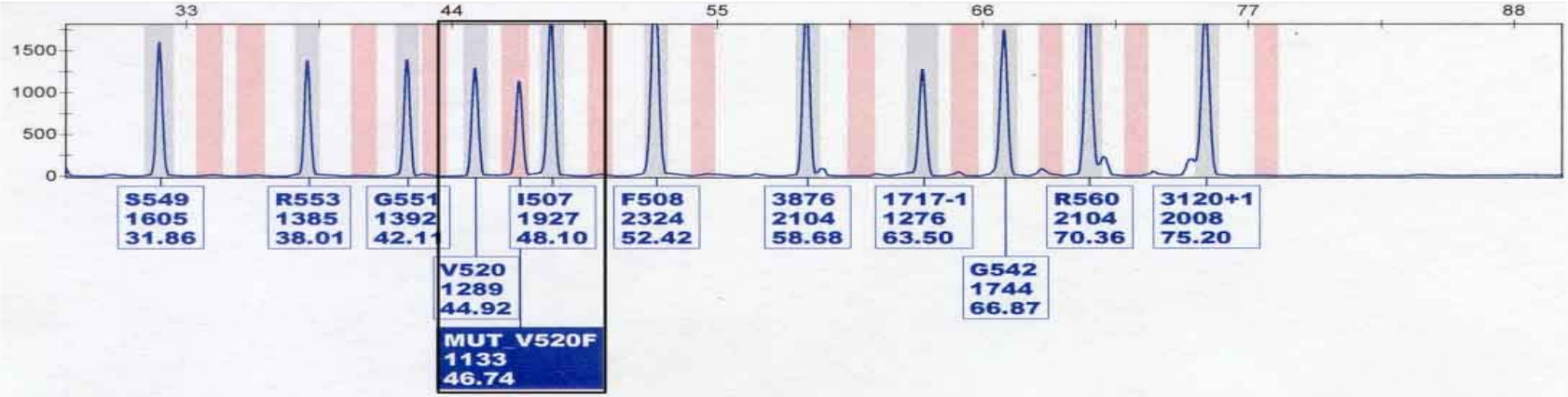
The third kind of error is **sporadic error**: a design flaw or implementation flaw such as when the sequence where your primer sits has an unexpected mutation and you get a false negative.

Evaluating a new <u>or changed</u> method for all of its mutations or analytes is important here – some commercially available controls and panels can be helpful here.

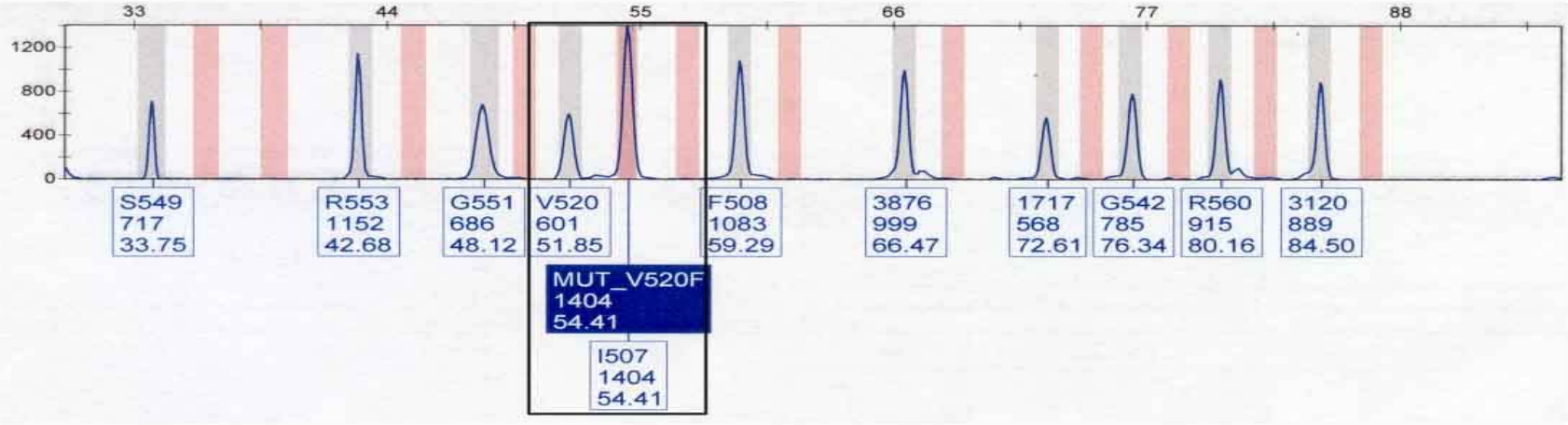
SeraCare
LIFE SCIENCES

A close encounter of the third kind…

V520F WGA on POP6 Polymer



V520F WGA on POP7 Polymer



SeraCare
LIFE SCIENCES

A quick review of error, continued

Total allowable error: what is it, and why do I care?

- A percentage of the difference in results that will trigger a medical intervention
- The %CV allowable under the CLIA regs (for some chemistry analytes)
- EP15 assumes if precision and accuracy are ***acceptable***, the assay's total analytical error is acceptable
- For additional info see: Krouwer JS. Setting performance goals and evaluating total analytical error for diagnostic assays. *Clin Chem.* 2002;48:919-927.

SeraCare
LIFE SCIENCES

No fair looking at the outcome and deciding "Looks okay to me…"

How to determine acceptable levels of precision, accuracy?

It's not as hard as it sounds:

- **Precision:** look at control performance (sd, %CV) for this or similar tests, and think about whether that's okay or you really need better.
- **Accuracy:** consider what the correlation coefficient (r^2) should be when you do your method comparison study.
- **Qualitative test:** decide what level of discrepancies between methods you can tolerate.

SeraCare
LIFE SCIENCES

Accuracy

Method comparison study

- Requires validated reference method, or closest thing you have
- Standards, controls, panels and/or proficiency test samples, plus some patient samples, using both assays, narrow time frame
- **Quantitative tests**: plot as difference and scatter plots
- **Qualitative tests:** plot on Contingency Tables
  - 2x2 table for each analyte

SeraCare
LIFE SCIENCES

2x2 Contingency Table, from EP12

Example 1a. 2 x 2 Contingency Table for Candidate Method vs Diagnostic Accuracy Criteria

		Diagnostic accuracy criteria: *H. pylori*		
		Positive	Negative	Total
Candidate Method	Positive	57	2	59
	Negative	4	39	43
	Total	61	41	102

Estimated Sensitivity = 100x(TP/TP+FN) = 100x(57/61) = 93.4%

Estimated Specificity = 100x(TN/FP+TN) = 100x(39/41) = 90.7%

When comparative method is not a reference method, use % agreement

SeraCare LIFE SCIENCES

Example from Cervista HPV 16/18 DNA

Table 1. Concordance with expected result (Cell Line Samples Only):

Enter results for samples S1a to S24a and S1b to S24b

		Cervista 16 18	
		+	-
Expected	+	16	0
Result	-	0	32

Number of indeterminate results:	0
Number of Samples in Agreement:	48
Total Number of Valid Samples:	48

formula = $\frac{\text{\# in agreement x 100}}{\text{\# valid samples}}$

48/48= **100%**

Criterion : >= 90%

Result: **Pass**

SeraCare LIFE SCIENCES

Does % agreement tell the whole story?

Method 1

	REF POS	REF NEG	
NEW POS	5	5	10
NEW NEG	5	85	90
	10	90	100

Method 2

	REF POS	REF NEG	
NEW POS	35	5	40
NEW NEG	5	55	60
	40	60	100

Overall % Agreement is 90.0% for each method, but are they equivalent?

PPA=50.0% (5/10)
[95% CI= 18.7%,81.3%]
NPA=94.4% (85/90)
[95% CI= 87.5%,98.2 %]

PPA=87.5% (35/40)
[95%CI=73.2%,95.8%]
NPA=91.7% (55/60)
[95% CI=81.6%,97.2%]

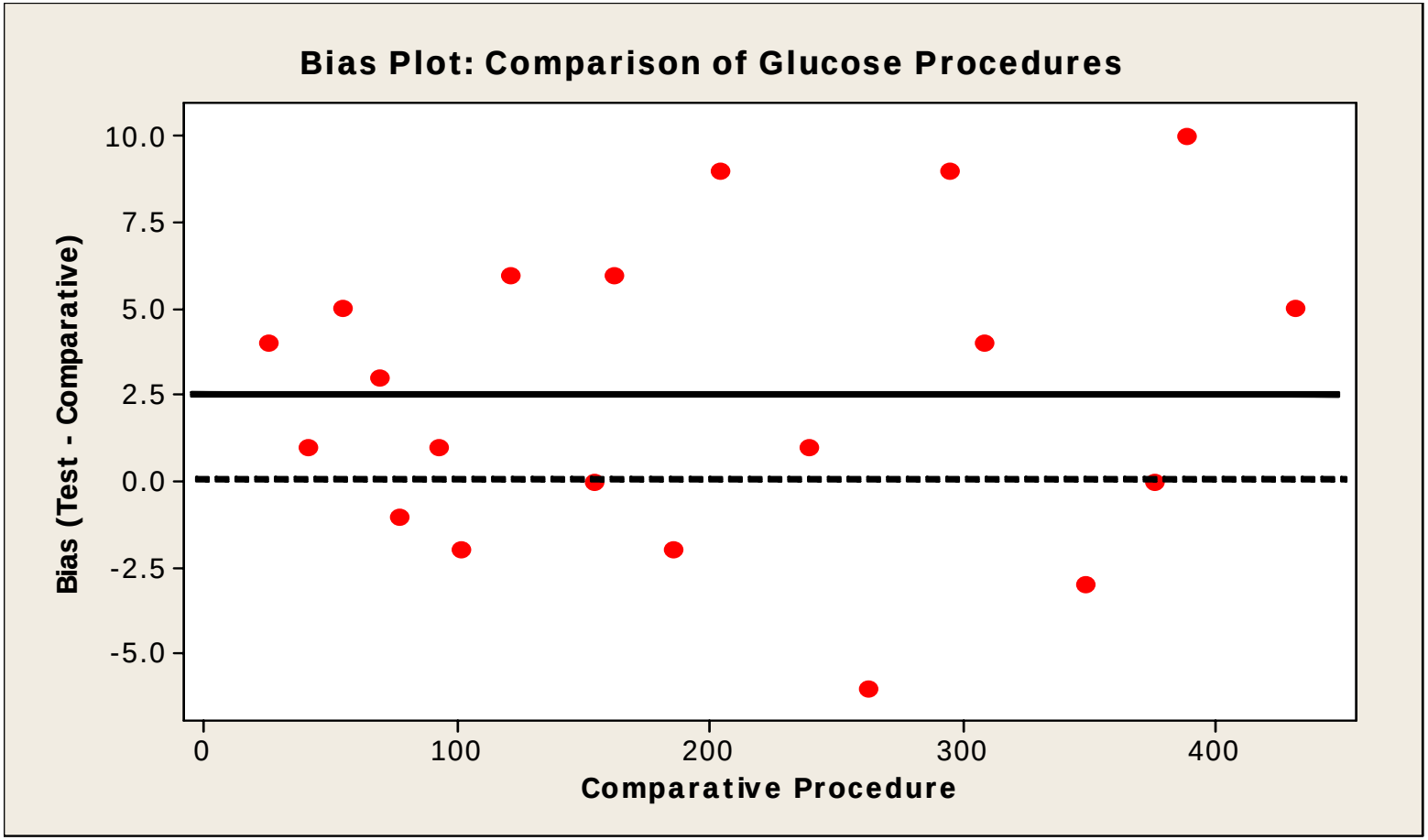
SeraCare
LIFE SCIENCES

Method Comparison Study: from EP15

- Minimum of 40 different patient specimens
  - Selected to cover a wide range of target concentrations
  - Fewer samples/broader range better than more samples/narrower range – check on reportable, reference ranges
- Duplicates of each sample preferable
  - Controls for operator issues, checks for outliers
- Multiple runs over several days
- Generate a “difference plot” and linear regression plot
  - Pearson Correlation Coefficient r shows random error
  - t-test shows random and constant errors

SeraCare
LIFE SCIENCES

Difference plot borrowed from EP15



$$\%b_i = 100 \bullet \left(\frac{\text{test procedure result}_i - \text{comparison procedure result}_i}{\text{comparison procedure result}_i} \right)$$

SeraCare LIFE SCIENCES

Linearity verification

Checks the useful analytical range

- EP15 recommends minimum 4 levels
- Preferable to use >5
- Make serial dilutions of a high sample, control or standard
  - panels for some analytes are commercially available
- Use clinically relevant sample matrix
- Run 3 replicates of each level
- Plot expected vs. observed (linear regression) and calculate regression line equation, correlation coefficient

SeraCare
LIFE SCIENCES

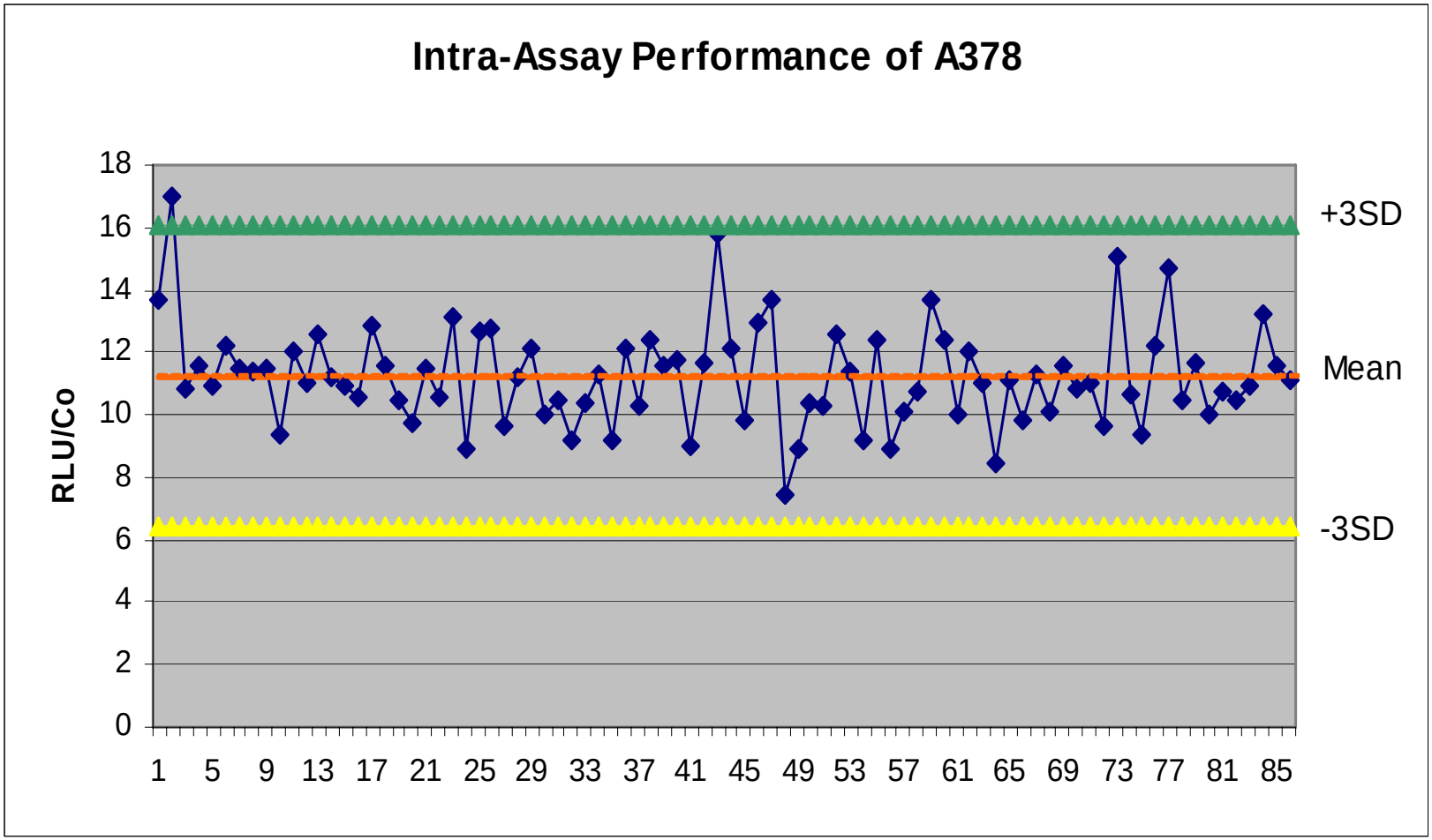
Precision verification I

Repeatability: results of multiple observations (tests) under identical conditions (often in one run)

- Estimates **random error** of a test
- Use **clinically relevant analyte levels**
  - low/high clinical decision points
- Use **clinically relevant sample matrix**
  - more than one?
- Typically involves ~20 replicates
  - calculate mean, sd, CV
- **QC samples customarily used**

SeraCare
LIFE SCIENCES

Repeatability – 86 replicates of SeraCare's HPV DNA control in a single run



SeraCare
LIFE SCIENCES

Precision verification II

Reproducibility: results of multiple observations (tests) under varied conditions

- Can allow estimation of random plus systematic error
- Multiple days, technicians, ***reagent lots***
- Same samples, analyte levels, number of observations, calculations – different variety of conditions
- Also uses control samples

Repeatability SDs should be <0.25 of Total Allowable Error (TE_a)

Reproducibility SDs should be <0.33 (TE_a)

SeraCare
LIFE SCIENCES

Analytical Studies

Analytical sensitivity: aka Limit of Detection (LOD)

- Lowest concentration of analyte that will be detected 95% of time with 95% confidence
- Minimally 20 replicates
- Narrow range of concentrations tested
- Last concentration with 19/20 positives is LOD estimate
- For multiplex assays, LOD must be determined for each target singly, as well as verified with other potential targets present.

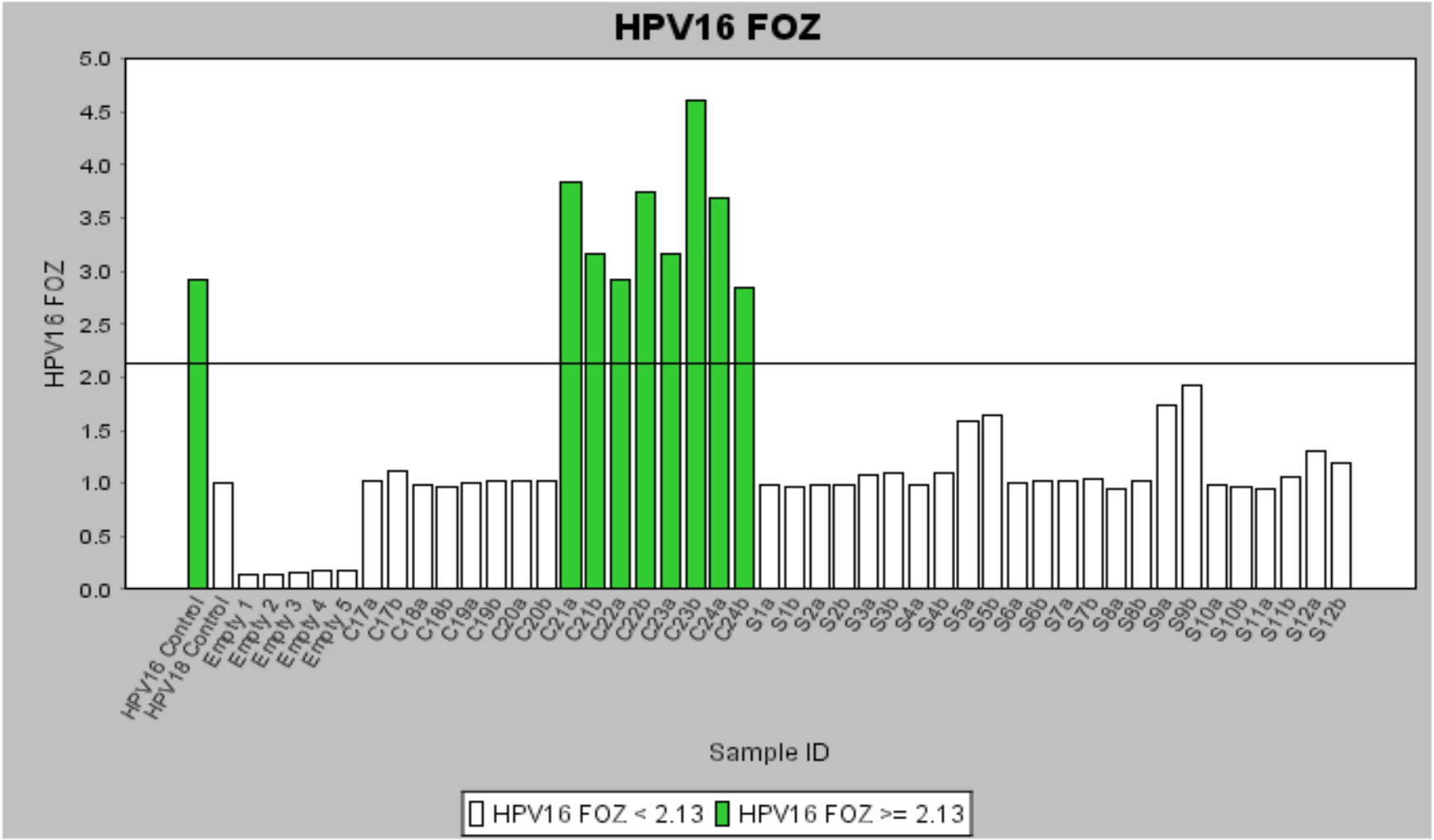
SeraCare
LIFE SCIENCES

Analytical Studies

Analytical specificity

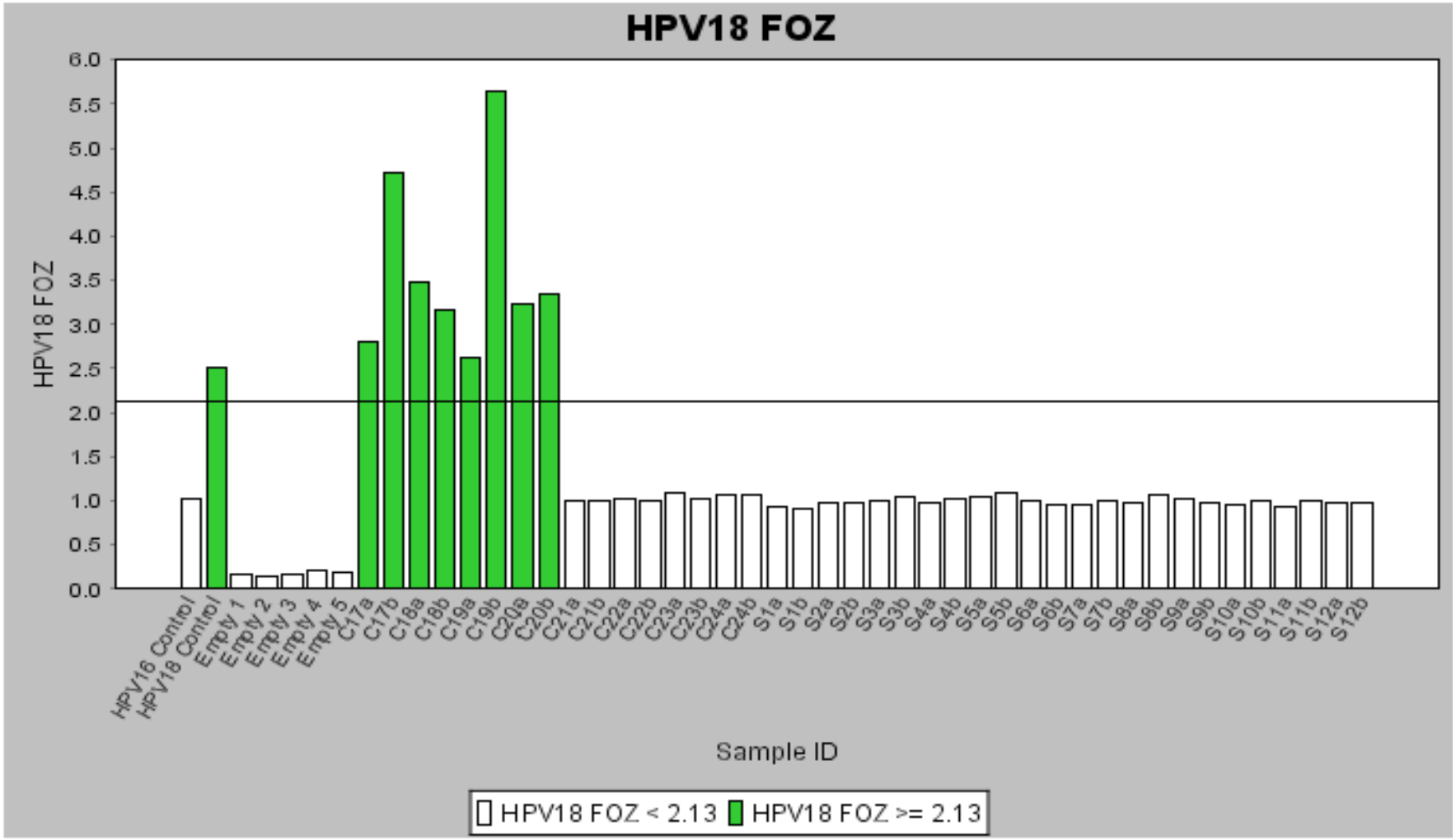
- Ability to detect desired target in the presence of other, closely related targets
- In the presence of interfering substances
- Performed using clinically relevant sample matrix
- For genotype assays, evaluate multiple alleles
- For ID assays, closely related species
- For multiplex assays, rule out cross-reactivity of detection reagents for other target in the assay

Cervista HPV 16 DNA Detection



SeraCare
LIFE SCIENCES

Cervista HPV 18 DNA Detection



SeraCare
LIFE SCIENCES

Other test parameters

Robustness: what virtues, glitches did you see during studies? Write those down!

Passes proficiency tests: PT samples tested in Method Comparison – how did they do?

Availability: more than one kit lot at a time? Consistent vendor responses?

Calibration plan: can you follow manufacturer's recommendations?

If a vendor generates your verification data…

- That's good, but the Lab Director still has responsibility for the quality of the results.
- Review the vendor's verification protocol with the rep, ask questions.
- Use tech rep availability to learn about technology and test method.
- Show the tech rep your FMEA, ask for his/her suggestions.
- In the vendor's protocol, are the main points described here (and in EP12 and EP15) covered?

SeraCare
LIFE SCIENCES

Planning a Quality Control Program

Quality control plan

Make one!!

- Is this new test among the most robust, or is it too soon to say?
- Use more QC for the tests you understand least well, or those you know well, but have less confidence in.
- Look again at your FMEA after verification. Should there be changes, now that you know more about the method?
- Where are the weak points in the test? That's where the QC should be the most intense.
- You can always decrease the amount of 'extra' QC as you become more confident with the test.

SeraCare
LIFE SCIENCES

Points to consider in formulating a QC plan

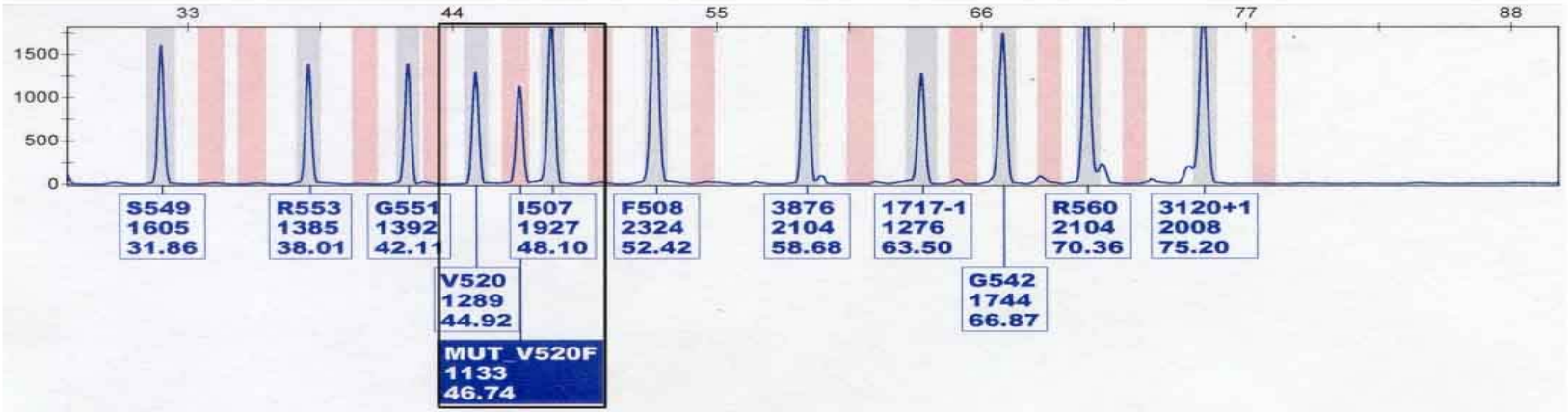
Every mutation or organism or compound or enzyme that you test for is a separate analyte.

- If you don't have a control for each one, how do you know that your test is detecting each one every time you run it?
- For your quant viral load tests, are you tracking the precision at 500 vs. the precision at 500,000? Do you think they are the same? Can you prove it?
- Do your controls monitor every step of the process?
- Yes, QC labor and materials cost – but what is the cost of a clinically significant error?
- Does proficiency testing take the place of QC?

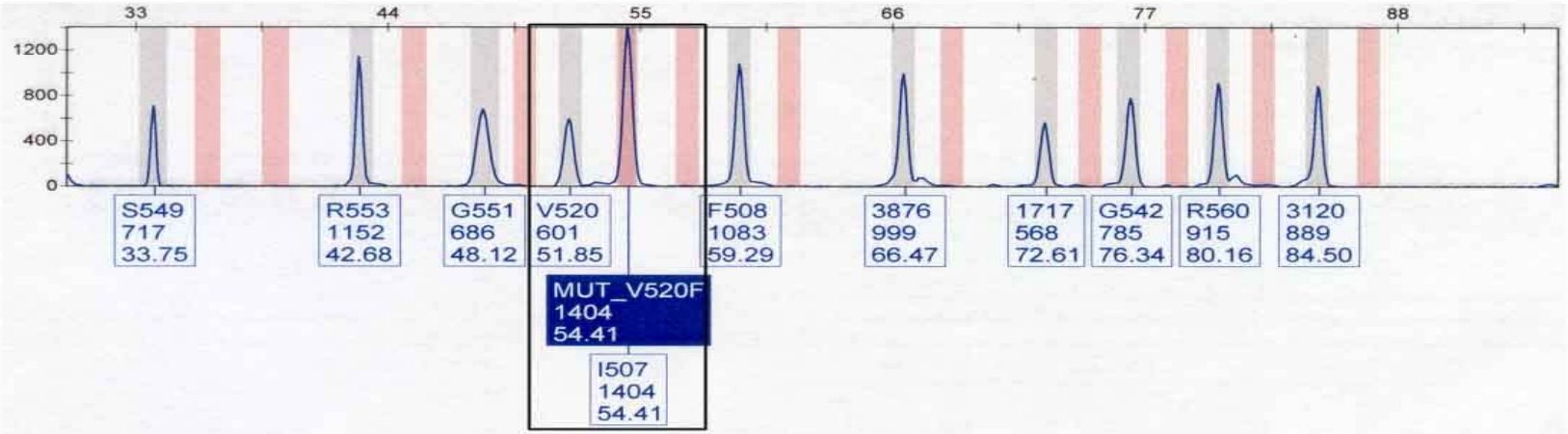
SeraCare
LIFE SCIENCES

With a simple polymer change…

V520F WGA on POP6 Polymer



V520F WGA on POP7 Polymer



SeraCare
LIFE SCIENCES

What QC monitoring can do, cannot do

QC can do this:

Detect **systematic error** (shift in the mean)

Detect **random error** (increase in the standard deviation)

Help you become familiar with and understand your test method (because you're consistently running a sample that you know the result of)

QC can't do this:

Detect sporadic error (design flaw)

Tell you exactly what went wrong (although maybe sometimes it can – if you control different assay steps separately)

SeraCare
LIFE SCIENCES

ARC Uses ACCURUN Low Positive HIV 1/2 Run Control to Monitor Tests, Accept Kit Lots

SERACARE POSITIVE EXTERNAL RUN CONTROL (Used by 5 Facilities in One Testing System)				
	Low to Mid-Range Positive External Run Control (PERC) to Meet NY State Regulations	**"Known positive" Sample for Qualification of New Assay Lots and/ or New Shipments**	**August 2009 S/CO for "Known Positive" Sample for Qualification of EIA Lot XXX**	**Immediate Actions Taken**
Facility 3	N/A- This facility does not test NY State samples.	Seracare PERC… Success= S/CO > 1.0	On 8/6/09, Seracare PERC Lot 113147- S/CO failed at 0.870. Repeat attempts failed, as well. Current EIA lot in production has acceptable PERC S/CO @ 2.08.	EIA LOT XXX was placed into quarantine in all 5 facilities, pending investigation by the assay manufacturer.
Facility 5	N/A- This facility does not test NY State samples.	A known positive sample is used (not Seracare PERC)… Success= S/CO > 1.0	First to receive EIA Lot XXX. Accepted based on S/CO values for 2 known positive samples @ 3.83 and 4.37. **(Usual mean: ~8)** On 8/6/09, this lot was qualified and ready for release once the previous lot was depleted.	Facility 5 has changed qualification process to include Seracare PERC, thereby standardizing the process for all 5 facilities.

SeraCare
LIFE SCIENCES

An ACCURUN (and QC) Success Story!

Lessons learned:

- ARC's QC program worked!
  - Labs followed SOPs
  - Reported failures at once to central communication facility
  - Early communication prevented use of less sensitive kit lot
  - Lot acceptance controls were standardized across facilities
- Kit manufacturer changed lot release criteria from qual to quant

SeraCare
LIFE SCIENCES

Role of Independent Controls in routine QC

- Can allow monitoring across kit lots, among test sites, over time.
- Can control the whole assay process, from extraction through detection, and supplement internal controls that are used post purification.
- Can allow continual comparison of site performance – multiple labs within an organization or across multiple sites of unrelated labs offering the same tests.
- Can allow comparison among different methods for the same analyte

SeraCare
LIFE SCIENCES

Thank you! And I hope we can answer your questions…

Pat Garrett, Ph.D., DABCC

SeraCare Life Sciences, Inc.

E-mail:

pgarrett@seracare.com

SeraCare
LIFE SCIENCES