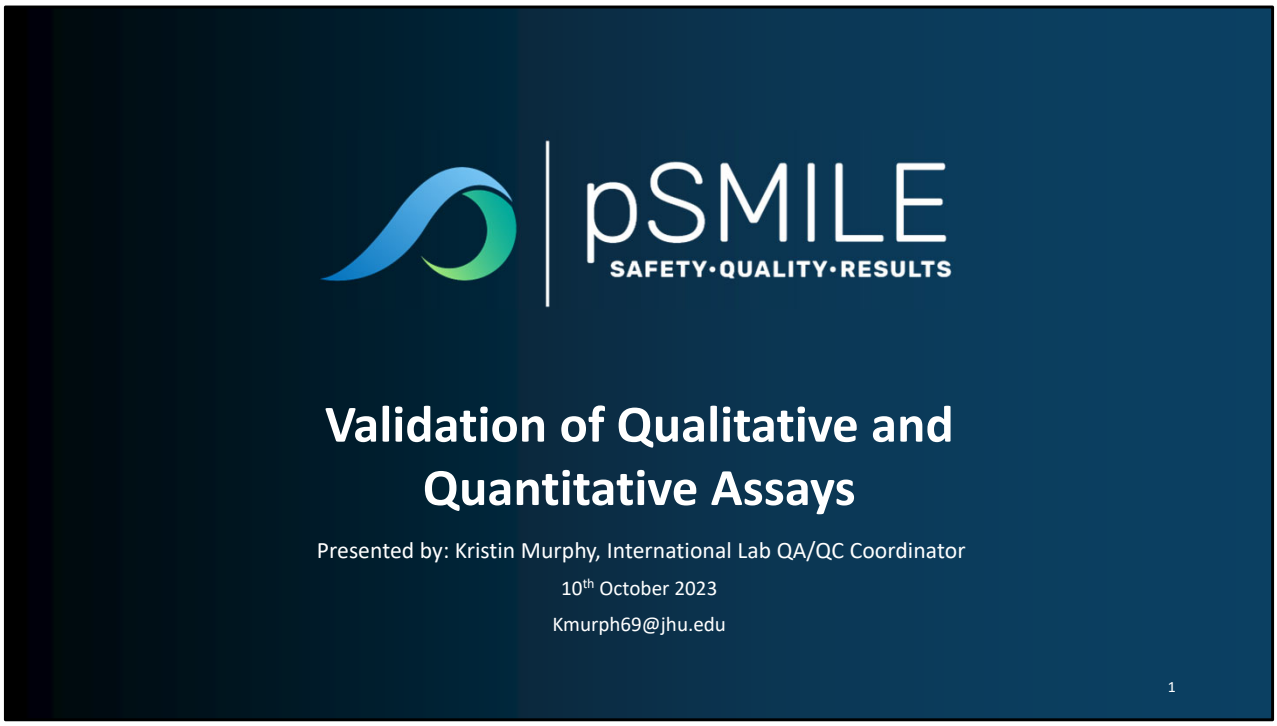

pSMILE
SAFETY·QUALITY·RESULTS
Validation of Qualitative and Quantitative Assays
Presented by: Kristin Murphy, International Lab QA/QC Coordinator
10th October 2023
Kmurph69@jhu.edu
1

This project has been funded in whole or in part with Federal funds from the Division of AIDS (DAIDS), National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services.
Contract Number: 75N93020C00001
Project Title: Patient Safety Monitoring in International Laboratories

Objectives

- Determine the essential components required for the validation of qualitative and quantitative assays
- Understand each component of validation including precision, accuracy, linearity, and reference ranges
- Work through practical examples of point-of-care, hematology, and chemistry validations
- Navigate the Resources sections of the pSMILE website for validation tools and templates

Agenda

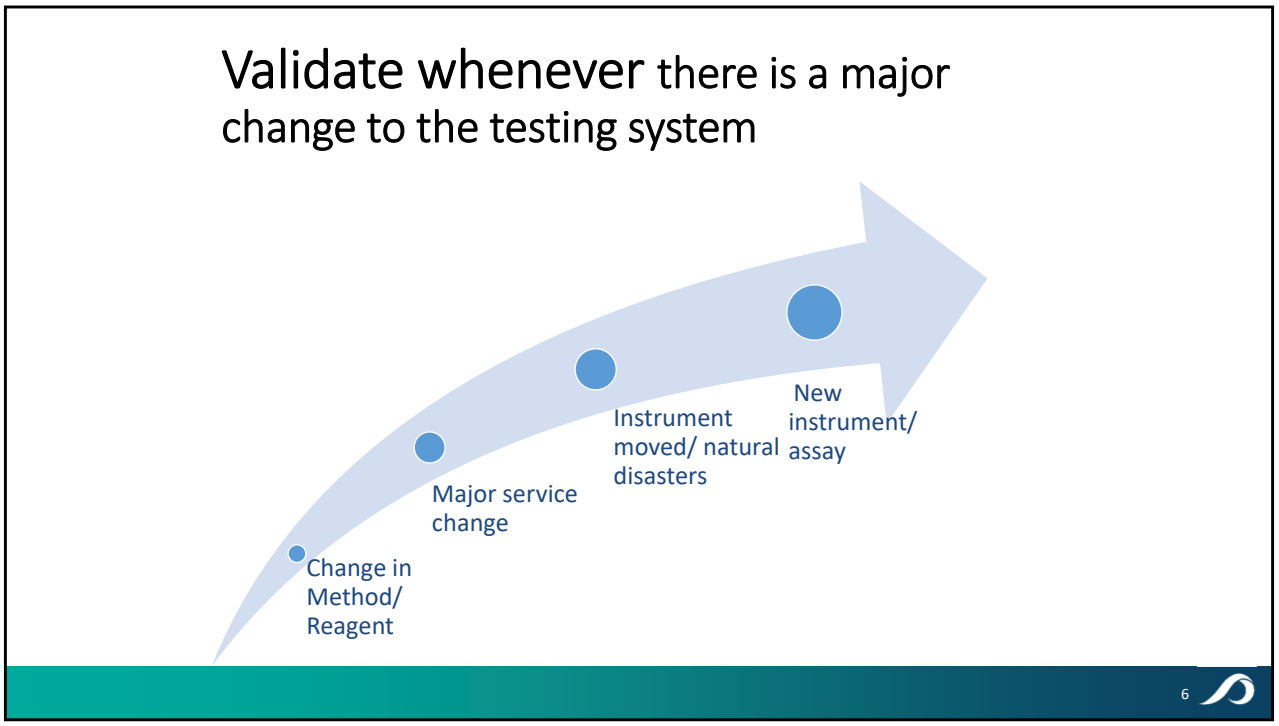
- Introduction/Precision
- Accuracy
- Linearity, AMR, CRR
- Reference Ranges
- Questions and comments

Main Elements of Validation

- Validation is the **verification** of:
  - Precision
  - Accuracy
  - Linearity
  - Analytical Measurement Range (AMR) & Clinical Reportable Range (CRR)
  - Sensitivity & Specificity
    - FDA-approved/ non FDA-approved
  - Reference Ranges

5

So, before we dive into precision, let's do a little overview of Validation. **Validation** is the verification that an analyzer or method is producing accurate and reproducible results. It is the process of verifying the manufacturer's claims for performance in order to ensure that the method will produce the expected results.



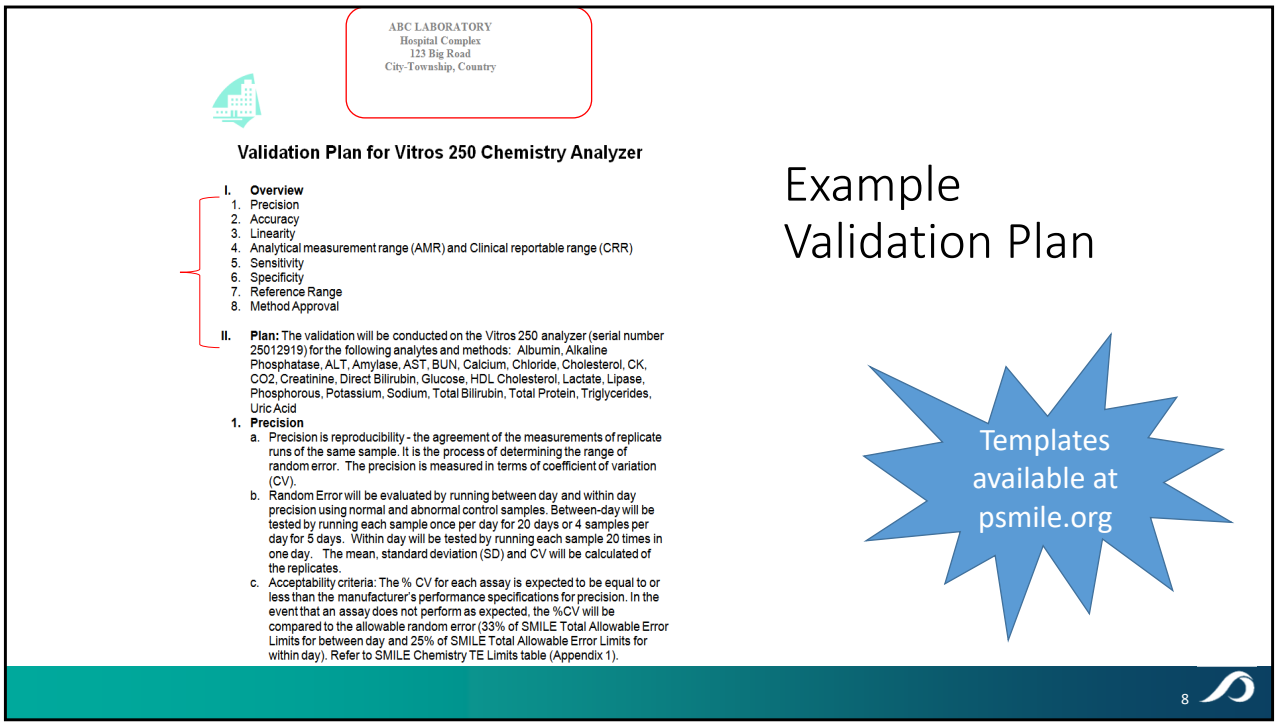
So when do we validate? We validate only following a **MAJOR** change.
1.) A change in reagent or method does **not** mean a simple lot change. It means that if the manufacturer has completely reformulated a reagent or changed methodology.
2.) Same goes for service. We're not talking about changing tubing or syringes—we're talking about a major change in the way an instrument operates.
3.) We also re-validate after an instrument has been moved, relocated, or if an instrument has undergone major trauma… with natural calamities such as earthquake.
4.) And of course—we always validate **before** using a new instrument or assay.

What is the first step?

- Make a plan for each instrument/method to be validated that includes:
  - What exactly is being validated: instrument, method, kit name, analytes, etc.
  - What type of samples will be used
  - Reference method used
  - Acceptability criteria for each type of testing

7

Before beginning any validation, the most important first step is to make a **PLAN.**



In SMILE website under resource, there is an available example of a validation plan that you can download and use as a template or a guide to help you create a customized validation plan for your laboratory. It's important to describe all of the steps that you plan to take in detail…writing this plan down will help you to anticipate problems and ensure that you have adequate staffing and materials available to complete the validation successfully.

Precision

Precision is reproducibility: the ability of a measurement to be consistently reproduced

So that's a very brief introduction to validation…let's move on to precision. **Precision** is always the first element of validation that should be completed. Precision is the ability to run the same sample, over and over again and get the exact same result. It is reproducibility.

How is precision measured?

- **Coefficient of variation (CV)**
  - A statistical measure of the dispersion of data points in a data series around the mean.

$$CV=\left(\frac{SD}{Mean}\right) x\ 100$$

  - Expressed as a percentage (%)
- Replication experiments are performed to estimate **imprecision** or **random error**

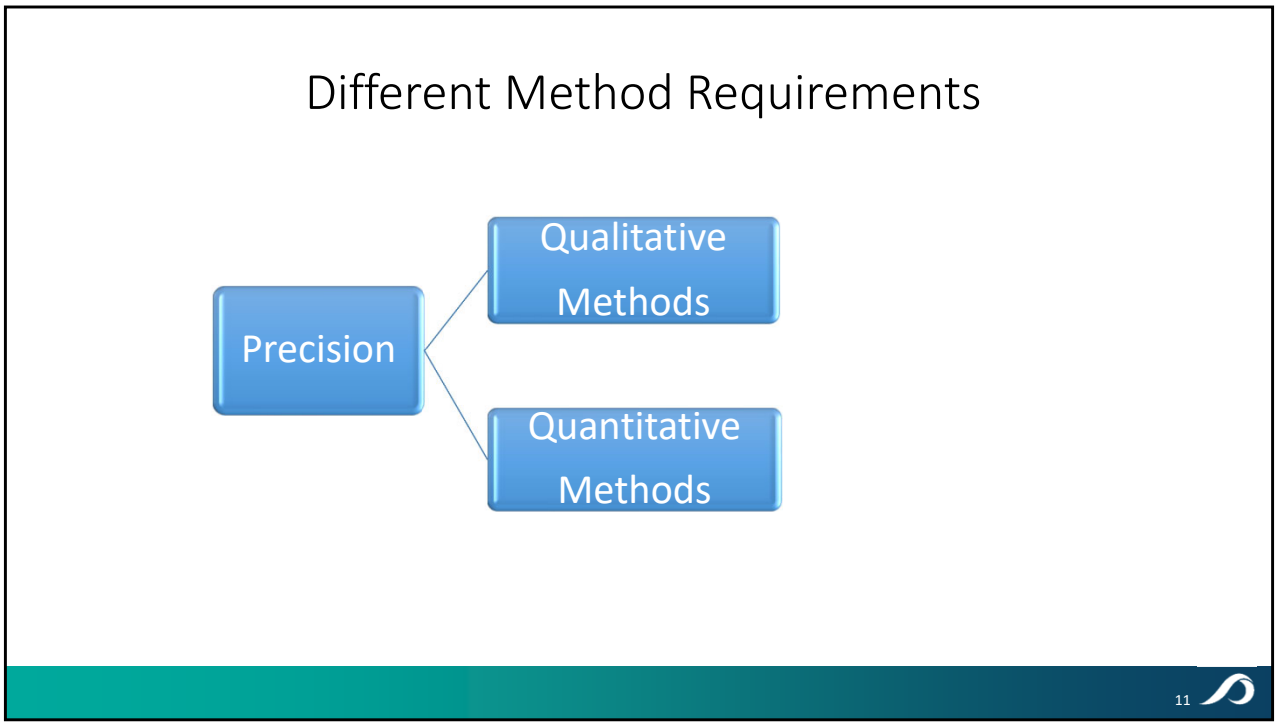
10

Precision is measured in terms on the CV or …
Stat calculation that measures the dispersion of datapoint around the mean

SD=Standard Deviation

Random errors in experimental measurements are caused by **unknown** and **unpredictable changes** in the experiment.

These changes may occur in the **measuring instruments** or in the **environmental conditions.**



We will be discussing precision measurements and evaluation in both qualitative and quantitative methods. The requirements and acceptability are slightly different.

When is Precision verification required for **qualitative** assays?

1. If the qualitative results are derived from a **quantitative value such as an OD.**

AND

2. Manufacturer's **package insert** describes precision specifications for the assay.

12

For DAIDS-funded laboratories, qualitative precision is only required to be verified in certain circumstances…..

Finding precision in package insert- **Qualitative** assays

TABLE II
ABBOTT PRISM HBsAg Assay Reproducibility

Panel Member or Control	Number of Replicates	Mean S/CO*	Intra-assay SD	Intra-assay %CV	Inter-assay[a] SD	Inter-assay[a] %CV
1	440	6.98	0.283	4.1	0.390	5.6
2	440	4.06	0.160	3.9	0.222	5.5
3	440	1.39	0.068	4.9	0.077	5.6
4	439[b]	8.86	0.513	5.8	0.596	6.7
5	438[c]	4.62	0.162	3.5	0.244	5.3
6	439[b]	1.37	0.078	5.7	0.083	6.1
7	440	0.34	0.036	10.6	0.039	11.6
Negative Control	439[b]	0.26	0.038	14.6	0.041	15.6

13

It is not always easy to find information in the package inserts. And if you do find it, it is often hard to decipher. Here is an example of a hepatitis package insert showing the manufacturer's precision results.

How to verify precision for **qualitative** assays

Short Term/ Within Run/Intra-assay Precision

- Samples: positive and negative controls
- Testing: Run each level of control 20 times on the same run, if possible

Long term/Between Run/Inter-assay Precision

- Samples: positive and negative controls
- Testing: Run each level of control at least once per day, but not more than 5 times per day, for a total of 20 runs

14

You could also use pooled patient samples for your positive and negative samples if the cost of using controls are too expensive.
Notice here you are using controls for this testing, just running the high and low or normal/abnormal 20 times.

Calculation and Acceptability – Qualitative Methods

1. Calculate the mean, SD, and CV of the **numerical result (OD)** for each level of control
2. Compare your CV to the manufacturer's CV
3. Lab CV should be ≤ manufacturer's CV

Qualitative Precision Example

ABBOTT murex

E
Murex HIV Ag/Ab Combination
REF 7G79-01 / 02
C14GE41GB
GE41/42
Read Highlighted Changes
Revised June 2009

Murex HIV Ag/Ab Combination

Table 4
Murex HIV Ag/Ab Combination - Assay Reproducibility

Specimen	Number of Assays	Number of Replicates	Mean Absorbance/ Cut-off ratio	Intra-assay %CV	Inter-assay %CV
Negative Control	4	10	0.266	8.7	11.3
HIV-1 Positive Control	4	10	8.287	4.3	4.7
QA01	4	10	3.672	4.6	7.3
QA02	4	10	4.696	5.6	12.9
QA03	4	10	3.006	3.9	4.2
QA04	4	10	1.663	6.8	9.2

Results- Acceptable or Not?

NOM_PRENOM	Abs	RESULTAT	DAT_TEST
	0.29	neg	05-Feb-09
	0.259	neg	05-Feb-09
	0.308	neg	05-Feb-09
	0.323	neg	05-Feb-09
	0.298	neg	05-Feb-09
	0.298	neg	05-Feb-09
	0.282	neg	05-Feb-09
	0.285	neg	05-Feb-09
	0.296	neg	05-Feb-09
	0.287	neg	05-Feb-09
	0.265	neg	05-Feb-09
	0.326	neg	05-Feb-09
	0.29	neg	05-Feb-09
	0.303	neg	05-Feb-09
	0.31	neg	05-Feb-09
	0.277	neg	05-Feb-09
	0.317	neg	05-Feb-09
	0.291	neg	05-Feb-09
	0.27	neg	05-Feb-09
	0.29	neg	05-Feb-09
Mean	**0.293**		
SD	**0.018**		
CV	**6.17%**		

NOM_PRENOM	Abs	RESULTAT	DAT_TEST
	3.628	POS	05-Feb-09
	3.588	POS	05-Feb-09
	3.59	POS	05-Feb-09
	3.547	POS	05-Feb-09
	3.498	POS	05-Feb-09
	3.533	POS	05-Feb-09
	3.595	POS	05-Feb-09
	3.68	POS	05-Feb-09
	3.528	POS	05-Feb-09
	3.644	POS	05-Feb-09
	3.485	POS	05-Feb-09
	3.558	POS	05-Feb-09
	3.521	POS	05-Feb-09
	3.659	POS	05-Feb-09
	3.6	POS	05-Feb-09
	3.642	POS	05-Feb-09
	3.651	POS	05-Feb-09
	3.54	POS	05-Feb-09
	3.566	POS	05-Feb-09
	3.654	POS	05-Feb-09
Mean	**3.585**		
SD	**0.058**		
CV	**1.63%**		

Precision for **quantitative** methods

- Precision verification is required for **all quantitative methods**

18

Let's move on to **precision for quantitative assays**.

Finding precision in package insert- **Quantitative** assays

SYNCHRON® System(s)
Chemistry Information Sheet

BUN
Urea Nitrogen
Kit Reorder # 442750

PRECISION

A properly operating SYNCHRON® System(s) should exhibit precision values less than or equal to the following:

TABLE 8 PRECISION VALUES

TYPE OF PRECISION	SAMPLE TYPE	1 SD		CHANGEOVER VALUE		% CV
		mg/dL	mmol/L	mg/dL	mmol/L	
Within-run	Serum/Plasma	2.0	0.71	66.7	24.0	3.0
	Urine	3.0	1.07	100.0	35.7	3.0
Total	Serum/Plasma	3.0	1.07	66.7	24.0	4.5
	Urine	4.5	1.61	100.0	35.7	4.5

Comparative performance data for a SYNCHRON LX® System evaluated using the NCCLS Proposed Guideline EP5-T2 appears in the table below.[15] Each laboratory should characterize their own instrument performance for comparison purposes.

For **Hematology instruments**- also see operator's manual

19

The first step is to determine what the manufacturer's stated precision specifications are. This can usually be found in the package insert. In case of Hematology instruments, you may need to look for this information in the operator's manual.

How to verify precision for **quantitative** assays

- Two Levels:
  - Normal/Abnormal
  - Hematology: Normal/High Abnormal
- Patient Samples *or* Quality Control

Hematology Controls-
Low level control is not recommended

20

Remember that for Hematology – low level control is **NOT** recommended,
Patient simples are not recommended for long term esp. hematology

How to verify precision for **quantitative** assays

Short Term/ Within Run/Intra-assay Precision

- Testing: run each level of control 20 times on the same run, if possible, or at a minimum within the same day

Long term/Between Run/Inter-assay Precision

- Testing: run each level of control at least once per day, not more than 5 times per day, for a total of 20 runs

21

Notice here you are using controls for this testing, just running the high and low or normal/abnormal 20 times.

Calculation and Acceptability – Quantitative Methods

1. Calculate mean, SD, and CV for each level using the 20 data points
2. Compare your CV to the manufacturer's CV
3. Lab CV should be ≤ manufacturer's CV
4. *If Lab CV > manufacturer's CV, compare to 25% or 33% of Total Allowable Error (TEa)*

22

If Short –Term/ Long Term precision is unacceptable, consult instrument manufacturer for assistance.

Total allowable error (TEa)

- Allowable error- the amount of error that can be tolerated without invalidating the medical usefulness of the analytic result
- pSMILE Recommendations for TEa are based on CLIA and are the same criteria used to evaluate EQA

23

SMILE recommendations for Tea are based on CLIA, endorsed by networks-these are the same criteria used to evaluate EQA.

pSMILE Recommendations for TEa

pSMILE Minimum Recommended Validation requirements for
Chemistry Total Allowable Error (TEa)

Analyte	pSMILE Total Error Limits (whichever is greater)		Precision	
	Percentage	Minimum detectable difference or absolute	Short Term 25% TE (1)	Long Term 33% TE (1)
Albumin	± 10% (1)	±0.2 g/dL 2.0 g/L	2.5%	3.3%
Alk. Phos	± 30% (1)	±5.0 U/L	7.5%	9.9%
ALT	± 20% (1)	±5.0 U/L	5.0%	6.6%
Amylase	± 30% (1)	±5.0 U/L	7.5%	9.9%
AST	± 20% (1)	±5.0 U/L	5.0%	6.6%
Bilirubin, Direct	± 20% (1)	± 0.4 mg/dL	5.0%	6.6%
Bilirubin, Total	± 20% (1)	± 0.4 mg/dL	5.0%	6.6%
Calcium	± 8% (2)	± 1.0 mg/dL 0.25 mmol/L	2.0%	2.64%
Chloride	± 5% (1)	± 2.0 mmol/L	1.25%	1.65%
Cholesterol	± 10% (1)	±3.0 mg/dL 0.08 mmol/L	2.5%	3.3%
CO2	± 20% (2)	±4.0 mmol/L	5.0%	6.6%
Creatinine	± 15% (1)	± 0.3 mg/dL 26.52 µmol/L	3.75%	4.95%

24

SMILE has compiled a list of the most commonly used analytes for protocol testing and their Total Allowable Error limits. In some cases, CLIA has not set a total allowable error limit for an analyte. In those cases SMILE has researched alternate sources (such as published studies on biological variation) and included those TeA limits in the table for easy reference.

For precision, 25% of TeA is allowable for short term (within-run) precision and 33% of TeA is allowable for long term (between day) precision. You should always attempt to meet the specifications given by the manufacturer—however this may not always be possible. If your precision data exceeds the manufacturer's specifications, it is acceptable to use these TeA limits as acceptability criteria.

Quantitative Precision Example

Package insert

Precision
Reproducibility was determined using human samples and controls in an internal protocol (within run: n = 21, between run: n = 10). The following results were obtained:

Sample	Within-run Mean U/L	Within-run Mean µkat/L	Within-run CV %	Between-run Mean U/L	Between-run Mean µkat/L	Between-run CV %
Human serum	58	0.97	1.8	58	0.97	3.2
Precitrol-N	32	0.53	2.1	32	0.53	3.2
Precitrol-A	124	2.07	1.1	124	2.07	1.8

Analytical sensitivity (lower detection limit)
4 U/L (0.07 µkat/L)
The detection limit represents the lowest measurable analyte level that can be distinguished from zero. It is calculated as the value lying three standard deviations above that of the lowest standard (standard 1 + 3 SD, within-run precision, n = 21).

	Normal	Abnormal
n1	37	206
n2	39	210
n3	38	210
n4	36	209
n5	38	206
n6	38	206
n7	42	205
n8	44	200
n9	41	204
n10	44	200
n11	41	204
n12	42	202
n13	41	199
n14	40	210
n15	38	203
n16	38	205
n17	36	204
n18	38	205
n19	37	211
n20	38	209

AST- Roche Cobas

Between Day

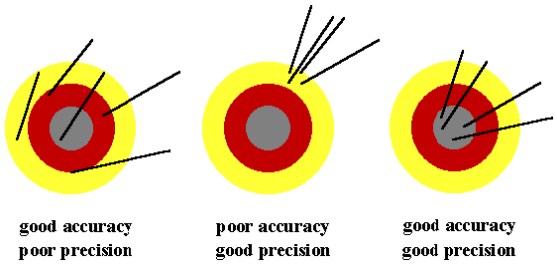
Normal CV: 6.2%
Abnormal CV: 1.7%

25

Normal: Acceptable. Lab CV is greater than package insert, but is within 33% TE.
Abnormal: Acceptable

Important Note

- **Precision experiments are performed to verify manufacturer's claims**



- **Remember to verify precision <u>first!</u>**

26

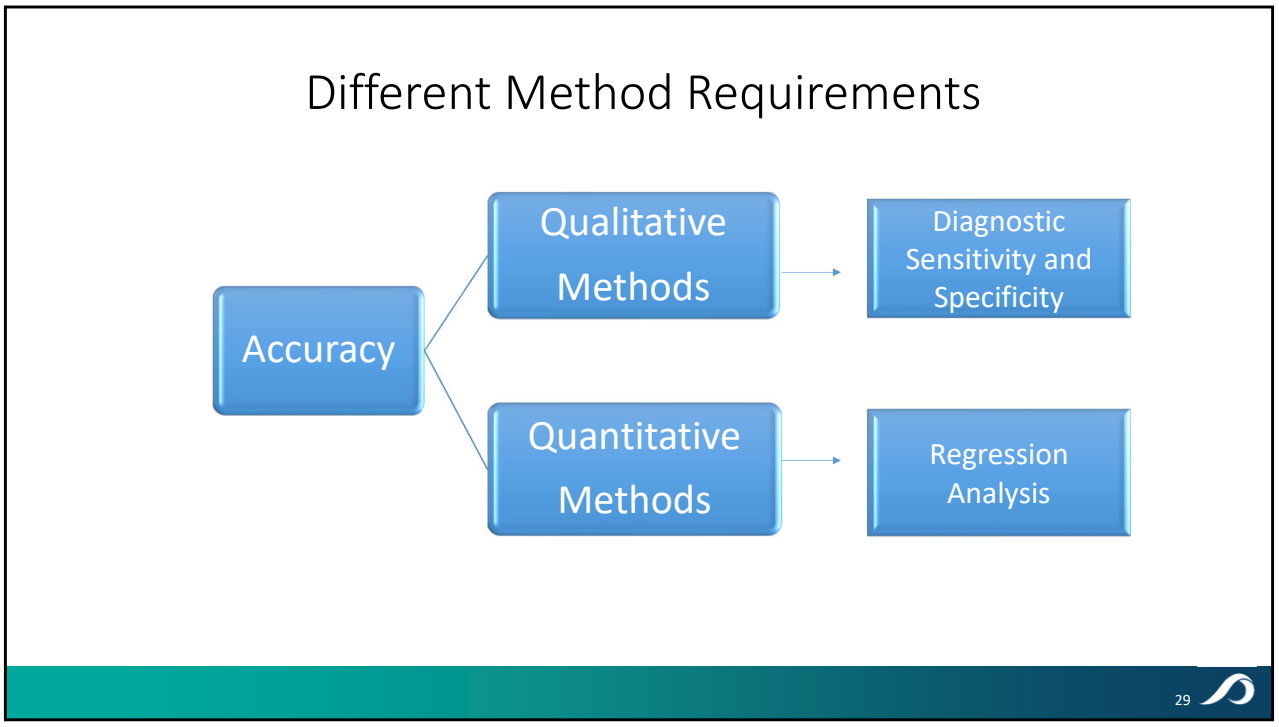
Always complete your precision verification FIRST. There is really no point in moving on to accuracy verification if you can't be confident that your instrument's precision is acceptable.


¿QUESTIONS?
27

Accuracy

Accuracy is the true value of a substance being measured.

Verification of accuracy is the process of determining that the test system is producing true, valid results.



We are going to talk about two broad types of Accuracy validation, these are Qualitative and Quantitative. We will look at some of the methods (kits and instruments) and samples of each, and then we'll look at how Qualitative Accuracy is evaluated by Diagnostic Sensitivity and Specificity, and Quantitative Accuracy is evaluated by Regression analysis and Error Index.

Determining Your **Qualitative** Reference Method

- The ideal reference method is a similar instrument/method
- Choose between:
  - An in-house reference method that has been previously validated and performing successfully on EQA
    - ➢ *Patient samples will be used*
  - EQA panels with known results
    - ➢ *EQA samples will be used*
  - Or a combination of both!

30

Sample Criteria

Sample Number

- 10 positive and 10 negative (or 10 of each expected result)

Other Considerations

- Rapid HIV tests: if using patient samples and method is non-FDA approved, may have additional confirmation requirements
- For urine hCG, manufacturer's stated cut-off limit should be considered

Diagnostic Sensitivity and Specificity

- The performance of qualitative tests is most commonly described in terms of **diagnostic sensitivity** and **specificity**
- Not to be confused with **analytic sensitivity** (lower limit of detection) and **analytic specificity** (interfering substances) that are another part of validation testing

Compiling the Results

- Once testing is complete develop a contingency table that compares the results of the qualitative test being validated with the results of the reference method

Abbott Murex anti-HCV (IDCP results)	Diagnostic Accuracy Criteria (Peer Results from CAP Surveys)		**Total**
	Positive	Negative	
Positive	***16 (True Positive)***	***0 (False Positive)***	***16 (TP+FP)***
Negative	***0 (False Negative)***	***19 (True Negative)***	***19 (FN+TN)***
Total	***16 (TP+FN)***	***19 (FP+TN)***	***35 (N)***

Calculations and Acceptability

- Use the table to calculate the following parameters and compare them to the manufacture's package insert

Diagnostic Sensitivity
100 x [TP/(TP+FN)]

Diagnostic Specificity
100 x [TN/(FP+TN)]

Positive Agreement
100 x [TP/(TP+FP)]

Negative Agreement
100 x [TN/(TN+FN)]

	Lab Result (%)	Expected Result	Acceptability
Sensitivity= 100 x [TP/(TP+FN)]	100%	99%	Acceptable
Specificity= 100 x [TN/(FP+TN)]	100%	99%	Acceptable
Positive Agreement (Positive Predictive Value) =100 x TP/(TP+FP)	100%	99%	Acceptable
Negative Agreement (Negative Predictive Value)= 100 x TN/(TN+FN)	100%	99%	Acceptable

Where to find the Manufacturer's Claims

PERFORMANCE CHARACTERISTICS

A multi-center clinical study was conducted to establish the performance of the QuickVue One-Step hCG-Urine test compared to results obtained from another commercially available hCG test. A quantitative method was used to resolve any discrepant results between the two test methods. In this multi-center field trial, 499 urine specimens, collected from patients presenting for pregnancy testing, were evaluated. A concordance of >99% was determined.

Urine Correlation
hCG Comparative Test

QuickVue hCG-Urine	+	–
+	252	0
–	0	247

Sensitivity: >99%
Specificity: >99%
Agreement: >99%

35

Determining Your **Quantitative** Reference Method

- The ideal reference method is a similar instrument/method
- Choose between:
  - An in-house reference method that has been previously validated and performing successfully on EQA
    - ➢ *Patient samples will be used*
  - EQA panels or commercial standards with known results
    - ➢ *EQA or standard samples will be used*
  - Or a combination of both!

36

Reference method is a little more complicated for quantive. Not all instuments and manufactures can be compared with eachother
Now we are going to branch out into Quantitative Accuracy. When we think of Quantitative tests, we think of instruments. How do we determine that results from an instrument are true and accurate? We recommend two main options for determining accuracy. Comparison to a previously validated instrument, or comparison with EQA. Comparison with a previously validated instrument is preferable, however EQA specimens may be necessary when there is no previous instrument, or in order to supply values in low and high ranges.

Sample Criteria

- Sample Number
  - At least 20 specimens, 40 is preferable
  - Tested in duplicate
- *For quantitative testing it is important that your accuracy specimens span the AMR of the instrument*

37

When selecting the 20 samples that you will use for accuracy, the ideal samples are patient samples. This will truly measure how your two methods compare on patient samples. QC, Calibrators and EQA may be subject to matrix effect and they are simulated samples. They are not ideal. However, it is often necessary to use a mix of patient and commercial samples because it is usually the only way to get samples that span the full range of the AMR.

Statistics used for Accuracy

Coefficient Correlation

- The correlation coefficient (R) must be >0.975

Slope

- The slope should be close to one

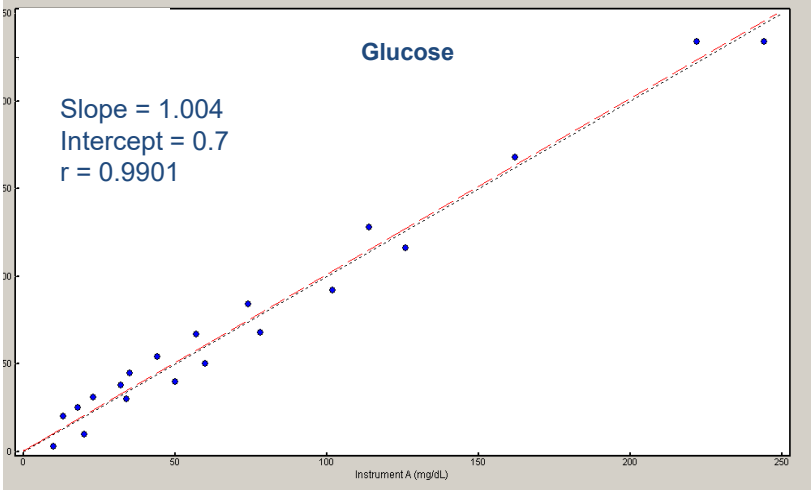
Intercept

- The intercept should be close to zero

38

Once we have run our 20 samples run in duplicate on both instruments, we need to analyze the data. Some of the statistics that we will use are correlation coefficient (r), slope and intercept. The R value measures the strength and the direction of a linear relationship between two variables. The slope represents the rise over run of the line drawn between the two data points (X and Y) and the Intercept is the point at which the straight line crosses the Y axis.
These statistics can be determined using any linear regression program (eg. Excel)

Accuracy Data Evaluation



39

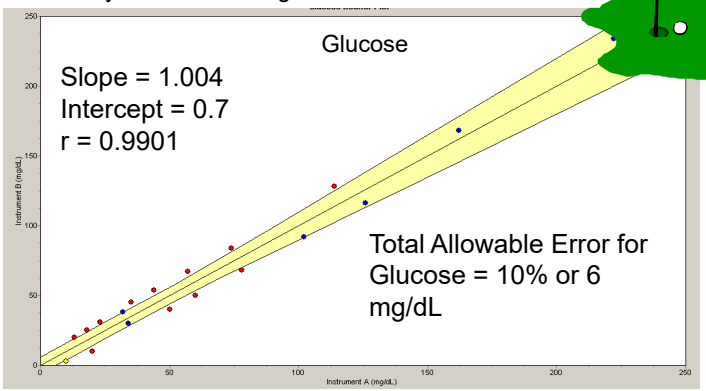
So, given those three typical statistical tools, what does it look like if we plot our accuracy data? Remember that the method being validated is always plotted on the Y axis. The reference method always goes on the X- axis.

So in this example, the correlation looks pretty good, right? Slope close to 1, Intercept close to Zero and the r is greater than 0.975.

There is no acceptable range, only ideals!

Slope should be “close” to 1.000
Intercept should be “close” to 0.0

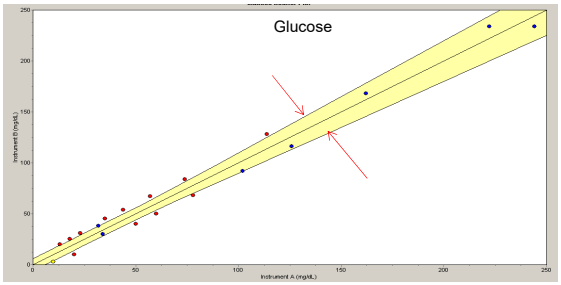
Obviously not close enough…



40

However, the problem with using only those three statistics is that none of the literature gives us absolute acceptability criteria. And we can see that when we evaluate the data using the Total Allowable Error, the validation fails.

Total Error (TEa) limit lines



- Error Index- Should be between -1.0 and 1.0

$$\frac{X - Y}{TEa}$$ For EACH X-Y pair

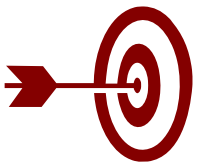
41

So, how should we evaluate accuracy data? By the use of Total Error Limit lines. In order to create these lines, the data analysis program calculates the "Error Index" for each X, Y pair. This is done by taking the X value, subtracting the Y value and dividing by Tea. In this way, we can evaluate if the difference between the two points is significant in relation to total allowable error. The error index for each X, Y pair should be between -1.0 and 1.0.

More statistics used for Accuracy

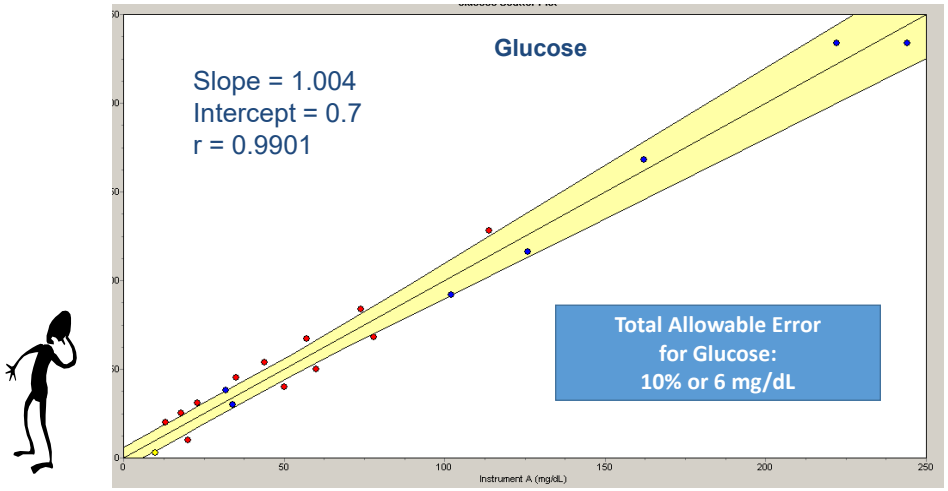
Error Index

- The “Error Index” measures the difference between the two methods as a ratio of the Total Allowable Error.
  - **Y** = New method
  - **X** = Comparison method



- Acceptability Criteria - The Error Index is measured for each X-Y pair.
  - The Error Index must fall within -1 and 1
  - For 95% of the specimens

Accuracy Data Evaluation with Total Allowable Error Limits



So you can see, that first example that looked pretty good using slope, intercept and r value—fails using total error limit lines.

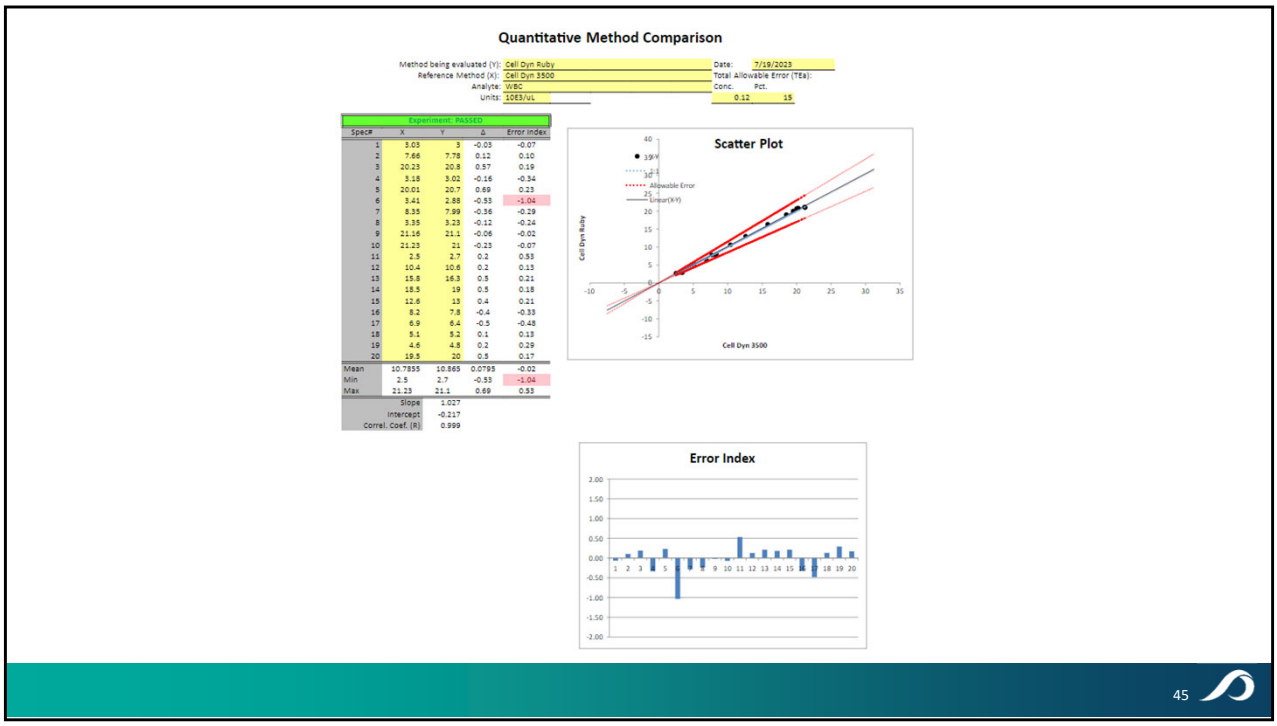
pSMILE Minimum Recommended Validation requirements for

Chemistry Total Allowable Error (TEa)

Analyte	pSMILE Total Error Limits (whichever is greater)		Precision	
	Percentage	Minimum detectable difference or absolute	Short Term 25% TE (1)	Long Term 33% TE (1)
Albumin	± 10% (1)	±0.2 g/dL 2.0 g/L	2.5%	3.3%
Alk. Phos	± 30% (1)	±5.0 U/L	7.5%	9.9%
ALT	± 20% (1)	±5.0 U/L	5.0%	6.6%
Amylase	± 30% (1)	±5.0 U/L	7.5%	9.9%
AST	± 20% (1)	±5.0 U/L	5.0%	6.6%
Bilirubin, Direct	± 20% (1)	± 0.4 mg/dL	5.0%	6.6%
Bilirubin, Total	± 20% (1)	± 0.4 mg/dL	5.0%	6.6%
Calcium	± 8% (2)	± 1.0 mg/dL 0.25 mmol/L	2.0%	2.64%
Chloride	± 5% (1)	± 2.0 mmol/L	1.25%	1.65%
Cholesterol	± 10% (1)	±3.0 mg/dL 0.08 mmol/L	2.5%	3.3%
CO2	± 20% (2)	±4.0 mmol/L	5.0%	6.6%
Creatinine	± 15% (1)	± 0.3 mg/dL 26.52 µmol/L	3.75%	4.95%

44

Let's talk for a minute about Tea and what it is. TeA is the total allowable error for an analyte. It's the amount (usually a percentage AND an absolute value) that has been determined to be an acceptable amount of error for an analyte. CLIA provides TeA limits for many analytes. For those that are commonly used at SMILE that don't have CLIA limits, we have looked to other resources to come up with SMILE TeA limits. The references for our TeA limits are all cited at the bottom of the table.



So how do you take all of this information and plug it into the SMILE evaluation tools? In order to plot your data you need several bits of information—You need the two methods you are comparing. Again, the reference method always goes on the X axis. You need the units of measure (making sure that both methods use the same units). You need the TeA for the method (both percent and absolute) and you need your raw data. Remember that ideally you should have data that was run in duplicate. What you will enter here is the mean of those duplicates.

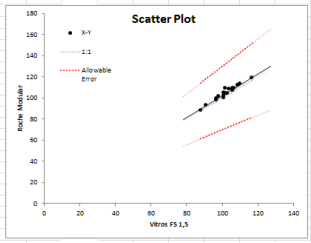
What do you notice about the data on this graph? Why is the line plotted so high up? And what statistic caused the validation to fail?
The R value is too low—which is an indication that the values don't span the AMR. Notice that none of the data points are near zero. There is no way to tell if these two instruments correlate any place except within the very tight range of around 90-120.

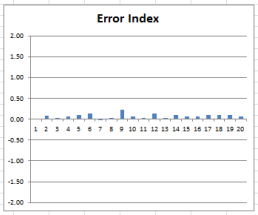
Quantitative Method Comparison

Method being evaluated (Y):	Roche Modular		Date:	
Reference Method (X):	Vitros FS 1,5		Total Allowable Error (TEa):	
Analyte:			Conc.	Pct.
Units:	U/L			5 30

Experiment: FAILED (See Note Below**)

Spec#	X	Y	Δ	Error Index
1	88	88	0	0.00
2	91	93	2	0.07
3	97	98	1	0.03
4	97	99	2	0.07
5	98	101	3	0.10
6	101	105	4	0.13
7	101	100	-1	-0.03
8	101	102	1	0.03
9	102	109	7	0.23
10	102	104	2	0.07
11	103	104	1	0.03
12	104	108	4	0.13
13	106	107	1	0.03
14	106	109	3	0.09
15	106	108	2	0.06
16	107	109	2	0.06
17	109	112	3	0.09
18	109	112	3	0.09
19	110	113	3	0.09
20	117	119	2	0.06
Mean	102.75	105	2.25	0.07
Min	88	88	-1	-0.03
Max	117	119	7	0.23
Slope	1.062			
Intercept	-4.143			
Correl. Coef. (R)	0.974	**Correlation Coefficient is too low. Troubleshoot Assay and re-analyse specimens		





46

How to Capture Accuracy Results

Document acceptability by filling in the table in your Validation Summary

- The correlation coefficient (R) must be >**0.975**
- The Error Index must be between **-1 and 1 for 95%** (19/20) of specimens
- Slope and intercept data should be reviewed for appropriateness

Analyze	Total Allowable Error	Correlation Coefficient (R)	Linear Regression Statistics	Linear Regression Statistics	Error Index Range	% of Error Indices -1.0 to 1.0	Acceptability
.		Expected >0.975	Slope	Intercept	Expected -1.0 to1.0	Expected ≥ 95%	
ALT	5.0 U/L or 20%	0.999	1.039	1.299	-0.14 to 1.20	90%	Unacceptable
AST	5.0 U/L or 20%	0.987	-1.009	0.083	-0.24 to 0.35	100%	Acceptable

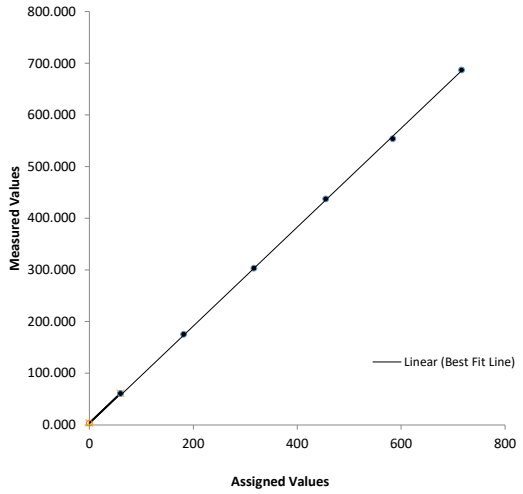
47

Now what we do with all of these statistics after we've analyzed our data. We always stress to our labs the importance of summarizing their data in a way that is easy to understand and is readily available for auditors. We have SMILE templates that call for each of the statistics that are most commonly used and asked for. All of these statistics are calculated using the SMILE evaluation tool. All you need to do is take the information from the spreadsheet and plug it into the template.

Linearity

A quantitative analytical method is said to be LINEAR when measured results from a series of sample solutions are directly proportional to the concentration or activity in the test specimens

LINEARITY



This means that a straight line can be used to characterize the relationship between measured results and the concentrations or activity levels

49

Note that there is a relationship between results means and the concentrations or levels of activity. There is a relationship, but the results are not necessarily equal to the concentrations. This is the difference between linearity and accuracy. With linearity, you can draw a line wherever you what on the graph and as long as it's straight, it's linear…but only a line at a 45 degree angle is accurate.

Sample Criteria

- At least 5 samples that cover the reportable range
- The values should be equidistant from each other
- Material:
  - Quality control
  - Calibrators
  - Commercial Linearity Standards

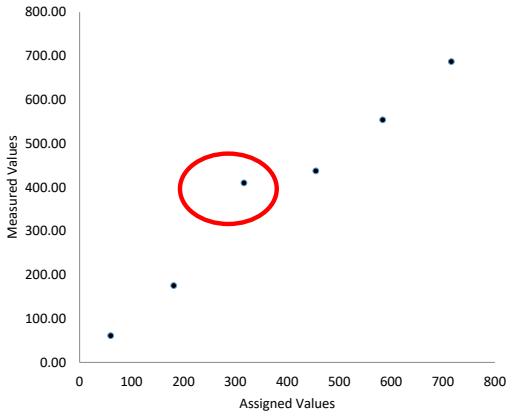
Contact pSMILE for Sources

50

What kinds of samples should we use for linearity?

Testing

- Test each sample in duplicate and average results
- Plot data immediately
  - pSMILE Linearity Worksheet
  - EP Evaluator
  - Any Regression Analysis Program
- Visually evaluate and correct any outliers!



51

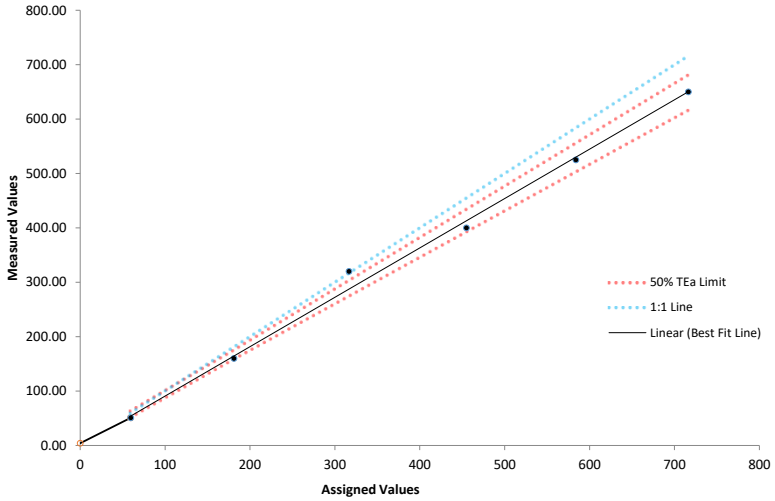
After we have our samples we are ready to test. Of course we want to make sure all our instrument maintenance is up to date, our internal QC is well within range and we haven't had any EQA failures prior to linearity testing.

You can use either a spreadsheet, EP, or other regression analysis program

Correct any underlying issues and rerun sample(s)

Data evaluation

Intercept -3.587
Slope 0.907
Correlation Coefficient 0.9997



52

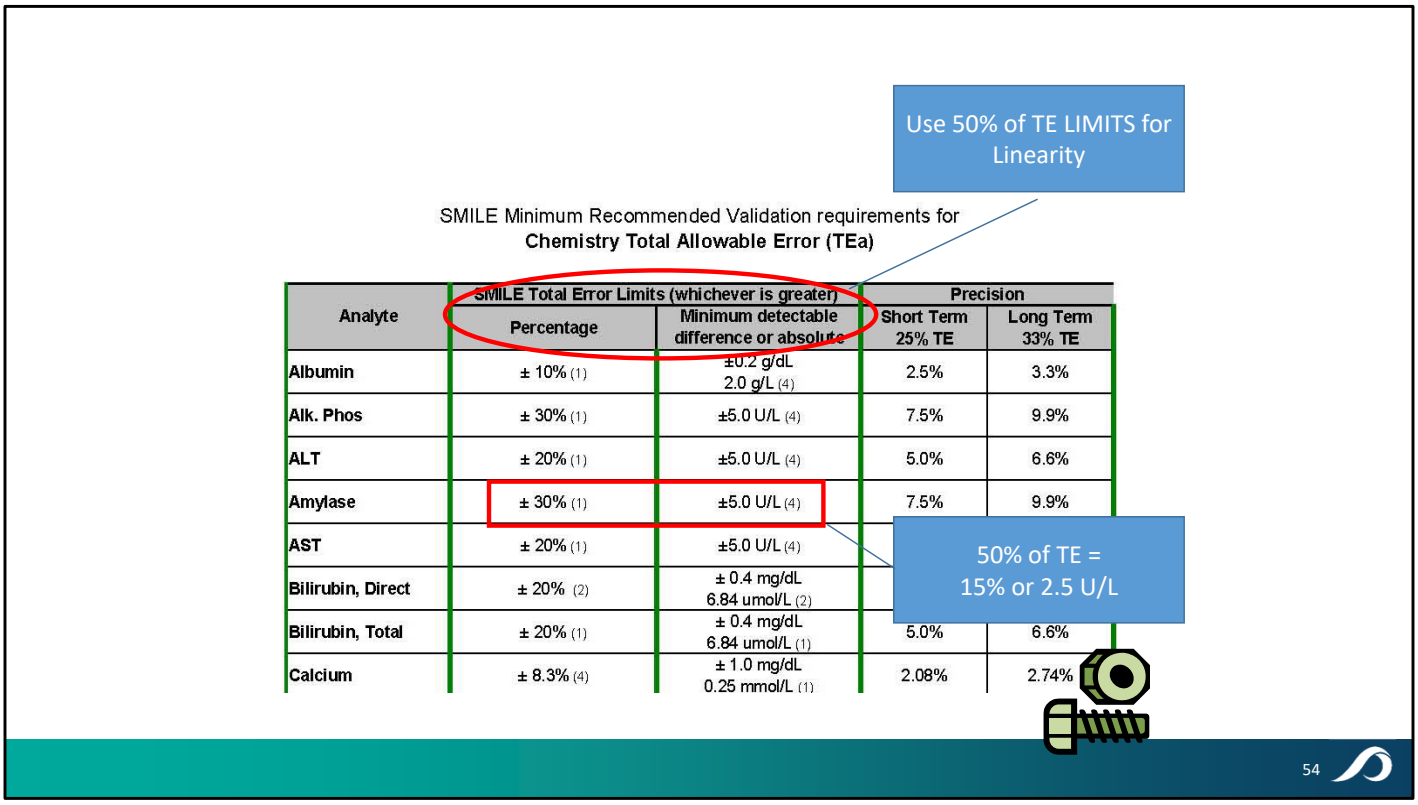
With the pSMILE spreadsheet you will get something that looks like this.

Acceptability Criteria

- The method is linear if the difference between the **predicted** Y and the **measured** Y is ***less than the allowable error*** for each specimen point
- The pSMILE Linearity spreadsheet and EP Evaluator will indicate "Pass" or "Fail" based on the above criteria

53

Other than visually, you want to evaluate your data by comparing the measured Y by the predicted Y.



Analyte	SMILE Total Error Limits (whichever is greater)		Precision	
	Percentage	Minimum detectable difference or absolute	Short Term 25% TE	Long Term 33% TE
Albumin	± 10% (1)	±0.2 g/dL 2.0 g/L (4)	2.5%	3.3%
Alk. Phos	± 30% (1)	±5.0 U/L (4)	7.5%	9.9%
ALT	± 20% (1)	±5.0 U/L (4)	5.0%	6.6%
Amylase	± 30% (1)	±5.0 U/L (4)	7.5%	9.9%
AST	± 20% (1)	±5.0 U/L (4)		
Bilirubin, Direct	± 20% (2)	± 0.4 mg/dL 6.84 umol/L (2)		
Bilirubin, Total	± 20% (1)	± 0.4 mg/dL 6.84 umol/L (1)	5.0%	6.6%
Calcium	± 8.3% (4)	± 1.0 mg/dL 0.25 mmol/L (1)	2.08%	2.74%

You remember the SMILE criteria table? You want to use 50% of the accuracy error limit to evaluate linearity

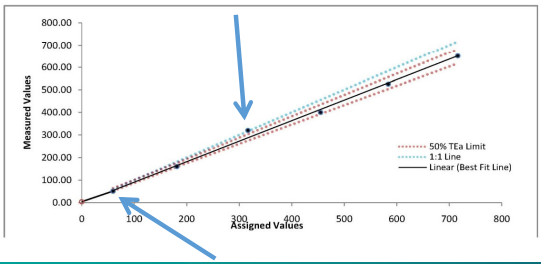
Data Evaluation using pSMILE Worksheet

Method Being Evaluated: Vitros 950/Glucose Oxidase - Color
Analyte: Glucose
Units: mg/dL
Date: 25-Mar-10
Total Allowable Error (TEa):
Conc.: 6 Percent: 10%
Enter values in ascending order
(lowest value data in top cells, highest value data in bottom cells)
Manufacturer's stated analytical measurement range:
Low: 10 High: 1000

Measured										
Rep 1	Rep 2	Rep 3	Avg	Assigned	Target	± Diff	% Diff	± Limit	% Limit	Grade
51	51		51.00	59.6	57.67	-6.67	-11.6%	6.00	5%	Unacceptable
160	160		160.00	181.2	167.15	-7.15	-4.3%	8.36	5%	Acceptable
320	320		320.00	316.6	289.06	30.94	10.7%	14.45	5%	Unacceptable
400	400		400.00	454.9	413.57	-13.57	-3.3%	20.68	5%	Acceptable
525	525		525.00	583.8	529.63	-4.63	-0.9%	26.48	5%	Acceptable
650	650		650.00	716.3	648.92	1.08	0.2%	32.45	5%	Acceptable

	Coefficients	Std Err
Intercept	4.010	14.327
Slope	0.900	0.032
Correl. (r)	0.9975	

Your Linearity Evaluation: LINEAR 400 to 650 PARTIAL RANGE
Verified AMR: 400 to 650 mg/dL



55

Here is the SMILE linearity worksheet. You simply enter the TEa in the Conc., & Percent blanks, then enter your manufacturer's stated linear range in the those blanks, then enter your data replicates into the first three yellow boxes, then enter your assigned values, and the worksheet does the rest. Notice Slope, Intercept and CC value. It evaluates where you are linear and gives you your AMR. It also plots your averages, draws a "best fit" line of your linearity, draws the acceptability (red) lines and the 1:1 (blue) line, or the line where Y would be equal to X. So you can see from this graph how two points are outside the error limit lines. So this data is only linear in the very narrow range of 400 to 600. You have to have at least three points to be linear. It would be quite easy if you could be linear just by drawing a line between two points.

How to Capture this in your Validation Report

3. Linearity and Reportable Range-refer to tab C

   i. Linearity

Analyte	Linear Regression Statistics		Allowable Systematic Error	Linearity Pass/Fail	Visual Evaluation	Acceptability
	Slope (Ideal=1.0)	Intercept (Ideal=0.0)	50% of CLIA	As evaluated by EP Evaluator		
ALT	0.970	0.282	10%	Pass	Linear	Acceptable
AST	0.931	0.42	10%	Pass	Linear	Acceptable
Albumin	1.018	0.36	5%	Pass	Linear	Acceptable

56

Then we just want to summarize our data in a chart something like this for the auditors. Notice the analytes in the first column. Then our regression statistics (slope and intercept), our allowable error from the SMILE table, whether our linearity passed or failed for each analyte according to the program that we used, how our data looks visually, and whether everything is acceptable or unacceptable. Also notice this “3. … refer to tab C. All the raw data, summaries, package inserts and any other information should be all organized in a binder with tabs. Linearity would be the third tab (C). This keeps everything organized for your auditors, PNL, etc.

Analytical Measurement Range (AMR)

The AMR is the range of analyte values that a method can directly measure on the specimen without any dilution, concentration, or other pretreatment not part of the usual assay process

AMR Validation

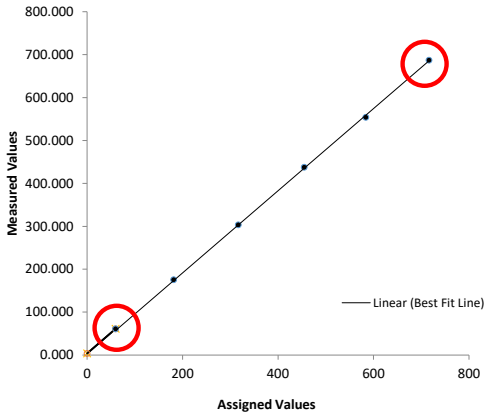
- **AMR validation** is the process of confirming that the assay system will ***correctly recover*** the concentration or activity of the analyte over the AMR
- The manufacturer defines the AMR
- It is the laboratory's responsibility to verify it

Sample Criteria

- Samples with an assigned or known value
  - Quality control
  - Calibrators
  - Commercial linearity standards

Sample Preparation

- Dilute the lowest sample to verify the low end of the AMR
- The high end of the AMR will only be as high as the highest sample



Working Example

SMILE Minimum Recommended Validation requirements for
Chemistry Total Allowable Error (TEa)

Analyte	SMILE Total Error Limits (whichever is greater)		Precision	
	Percentage	Minimum detectable difference or absolute	Short Term 25% TE	Long Term 33% TE
Albumin	± 10% (1)	±0.2 g/dL 2.0 g/L (4)	2.5%	3.3%
Alk. Phos	± 30% (1)	±5.0 U/L (4)	7.5%	9.9%
ALT	± 20% (1)	±5.0 U/L (4)	5.0%	6.6%
Amylase	± 30% (1)	±5.0 U/L (4)	7.5%	9.9%
AST	± 20% (1)	±5.0 U/L (4)	5.0%	6.6%
Bilirubin, Direct	± 20% (2)	± 0.4 mg/dL 6.84 umol/L (2)	5.0%	6.6%
Bilirubin, Total	± 20% (1)	± 0.4 mg/dL 6.84 umol/L (1)	5.0%	6.6%
Calcium	± 8.3% (4)	± 1.0 mg/dL 0.25 mmol/L (1)	2.08%	2.74%

- Total Bilirubin
  - Manufacturer AMR 0-25 mg/dL
- Allowable Error:
  - 20% or 0.4 mg/dL

61

Let's take a look at an example. Need to use Tea table to determine total error limits. Let look at bilirubin. We are told the manufacturers AMR is 0-25 mg/dL. The allowable error for bili is 20% or 0.4

Lower Limit Verification

Manufacturer's AMR: 0 - 25

- Need to verify **0 mg/dL**
  - (Lower Limit AMR)
- TE is 20% or **0.4 mg/dL**
  - (Whichever is greater)
- Need a standard within an assigned value from:

0 – 0.4 mg/dL

62

In this case use 0.4

Lower Limit Verification

Manufacturer's AMR: 0 - 25

- Bilirubin Lowest Standard Available
  - Assigned Value: 0.6 mg/dL
  - Subtract Total Error:

0.6 – 0.4 = 0.2 mg/dL

If you use this standard **without dilution**, this would be the lowest limit that could be accepted (after verification).

- If possible, **dilute the standard** to get within TE of the Lower Limit AMR

Lower Limit Verification

Manufacturer's AMR: 0 - 25

- Dilute Standard 1:2
  - 0.6/2 = 0.3 mg/dL Assigned Value
- Determine the Acceptable Criteria:
  - 0.3 ± 0.4 TE = 0–0.7 mg/dL
- Test the Standard:
  - Example Test Result = 0.40 mg/dL
- Evaluate Acceptability
  - 0.40 is within 0-0.7

Acceptable! Verified lower limit AMR is 0 mg/dL

64

Upper Limit Verification

Manufacturer's AMR: 0 - 25

- Need to verify **25 mg/dL**
  - (Upper Limit AMR)
- TE is **20%** or 0.4 mg/dL
  - (Whichever is greater)
- Need a standard within an assigned value from: **20 - 25 mg/dL**

25 x 0.2 = 5.0 mg/dL

25 – 5 = 20 mg/dL

65

In this case use 20% because it is greater

Upper Limit Verification

Manufacturer's AMR: 0 - 25

- Bilirubin Highest Standard Assigned Value:
  - 22.5 mg/dL
- Determine the acceptable Criteria:
  - 22.5 ± 20% TE =

22.5 x 0.2 = 4.5 mg/dL

22.5 ± 4.5 =

18 – 27 mg/dL

Acceptable Limits for 22.5 Standard: 18 - 25

Your acceptability limit shouldn't exceed the manufacturer's limit!

66

Upper Limit Verification

Manufacturer's AMR: 0 - 25

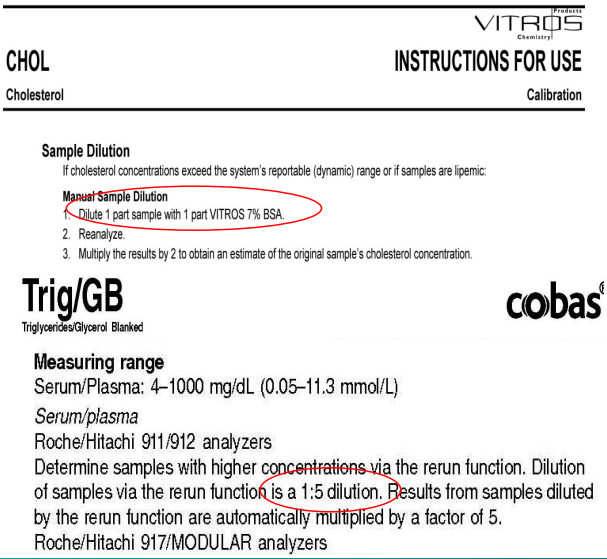
- Test the Standard:
  - Example Test Result = 21.0 mg/dL
- Evaluate Acceptability
  - 21.0 is within 18 – 25 mg/dL

Acceptable! Verified Upper limit AMR is 25 mg/dL

Clinical Reportable Range (CRR)

The CRR is the range of analyte values that a method can report as a quantitative result allowing for specimen dilution, concentration or other pretreatment used to extend the direct AMR

CLINICAL REPORTABLE RANGE



VITROS Products Chemistry

CHOL — **INSTRUCTIONS FOR USE**

Cholesterol — Calibration

Sample Dilution

If cholesterol concentrations exceed the system's reportable (dynamic) range or if samples are lipemic:

Manual Sample Dilution

1. Dilute 1 part sample with 1 part VITROS 7% BSA.
2. Reanalyze.
3. Multiply the results by 2 to obtain an estimate of the original sample's cholesterol concentration.

Trig/GB

Triglycerides/Glycerol Blanked

cobas®

Measuring range

Serum/Plasma: 4–1000 mg/dL (0.05–11.3 mmol/L)

Serum/plasma

Roche/Hitachi 911/912 analyzers

Determine samples with higher concentrations via the rerun function. Dilution of samples via the rerun function is a 1:5 dilution. Results from samples diluted by the rerun function are automatically multiplied by a factor of 5.

Roche/Hitachi 917/MODULAR analyzers

- How to determine what is appropriate:
  - Manufacturer's Recommendations
  - Literature References
  - Clinical Significance

69

Determining a CRR

- The laboratory should establish a CRR that covers a range inclusive of Grade 4 Adverse Events on the DAIDS Toxicity Table **without exceeding** the manufacturer's recommendations for dilution.

LABORATORY				
PARAMETER	**GRADE 1 MILD**	**GRADE 2 MODERATE**	**GRADE 3 SEVERE**	**GRADE 4 POTENTIALLY LIFE-THREATENING**
CHEMISTRIES	***Standard International Units are listed in italics***			
Bilirubin (Total)				
Adult and Pediatric > 14 days	1.1 – 1.5 x ULN	1.6 – 2.5 x ULN	2.6 – 5.0 x ULN	> 5.0 x ULN
Infant*†, ≤ 14 days (non-hemolytic)	NA	20.0 – 25.0 mg/dL *342 – 428 µmol/L*	25.1 – 30.0 mg/dL *429 – 513 µmol/L*	> 30.0 mg/dL *> 513.0 µmol/L*
Infant*†, ≤ 14 days (hemolytic)	NA	NA	20.0 – 25.0 mg/dL *342 – 428 µmol/L*	> 25.0 mg/dL *> 428 µmol/L*

How to Capture this in your Validation Report

4. Analytical Measurement Range (AMR) and Clinical Reportable Range (CRR)- refer to tab D

Analyte	Mfg's AMR	Low Value Verified	High Value Verified	Reportable Range	Dilutions	CRR	DAIDS Toxicity Grade 4
ALT	5-700 U/L	2.5	770	5-700	1:10	5-7000	>381
Total Bilirubin	0 – 25 mg/dL	0	25	0 - 25	1:10	0-250	>125

71

Analytical Sensitivity and Specificity

Analytical Sensitivity *is the lowest concentration of an analyte that can be measured (also called the Lower Limit of Detection).*

Analytical Specificity *is the determination of the effect of interfering substances*

Analytical Sensitivity and Specificity

- Unmodified/FDA approved method:
  - Refer to test package insert
- Modified/non FDA approved method:
  - The laboratory must establish the lowest concentration that the method can accurately measure that is distinguishable from zero
  - The laboratory must determine the effect of interfering substances
  - Consult with the Networks for requirements and recommendations

Reference Ranges

The range of test values expected for a designated population where 95% of the individuals are presumed to be healthy (or normal)

The last topic for today is reference ranges.

We always make the validation of reference ranges as the last part, after precision, accuracy and linearity.

Reference ranges are often incorrectly called "normal range". In fact, as shown in this picture, there is no such thing as "normal". This also illustrates why it is so important to establish reference ranges based on their own local population.

How do you validate reference ranges?

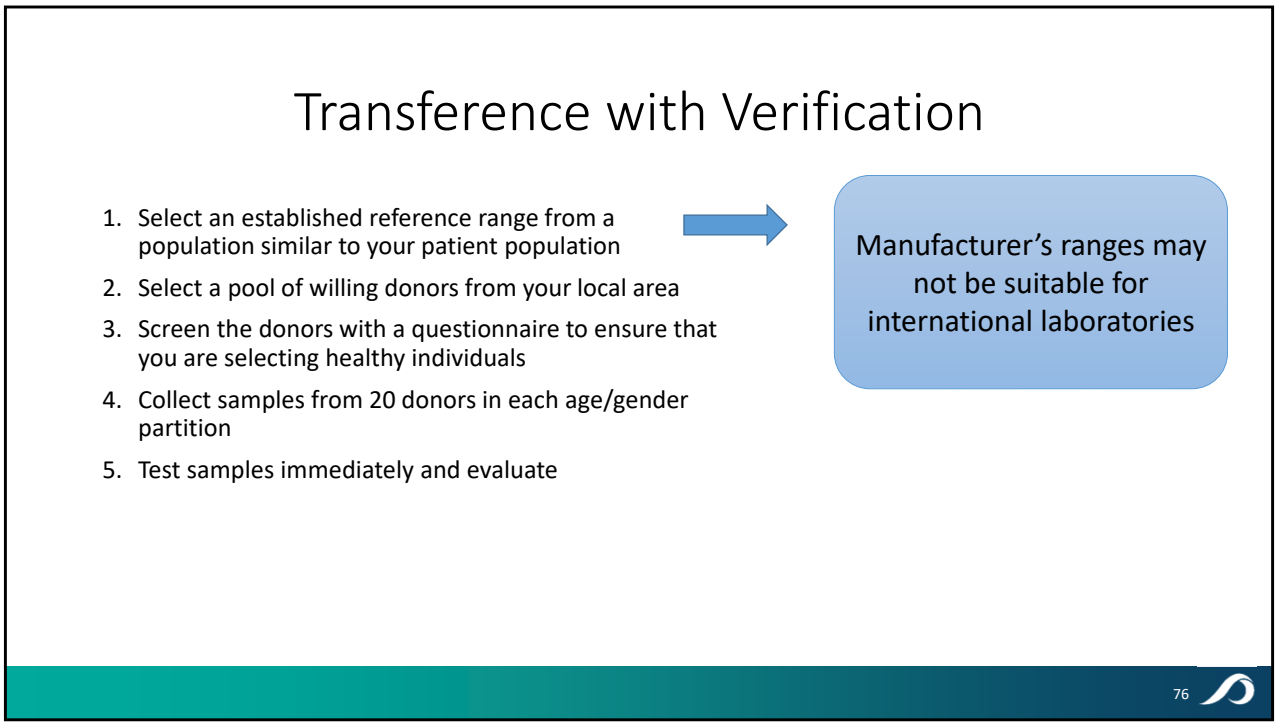
1. Transference of reference ranges (with verification)
2. Establishment of reference ranges
3. Transference of reference ranges (without verification)

75

Reference ranges must be validated for every new test method/instrument put into use in a laboratory. Without local reference ranges, there is no way to make appropriate medical decisions.
Reference ranges should be done as the final step in your validation. You cannot begin to validation reference ranges until all other aspects of the instrument have been validated.

Reference ranges should be done as a final step in validation. You can not begin to validate the ranks until all other aspects of the instrument have been validated.

We will not talk about transfer without verification ranges today. Only this method should be used for pediatric ranges.

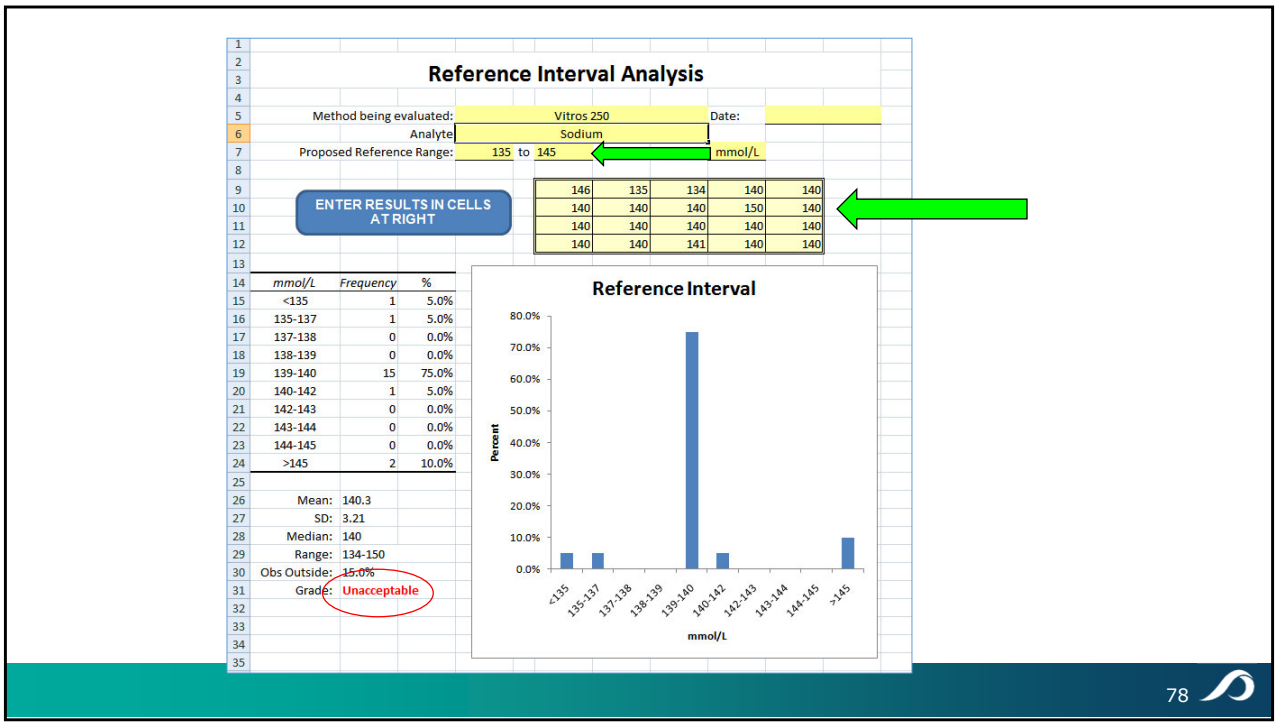


1. You must carefully select a range with similar patient demographics and similar geographic and nutritional conditions. Possible sources include—published reference ranges, a neighboring lab's reference ranges. Psmile.org contains a library of ref ranges from our labs that can be shared. It's important to select a range with similar geographics and demographics to your own population.
2. It is also acceptable to use you own, previously established reference range—you may have a range that has been in use at your institution. It is acceptable to select this as the range you will attempt to verify.
3. Also important to note that you don't have to choose only one "set" of reference ranges to verify. You can select from a variety of sources.
4. Willing donors are often your laboratory staff—try offering incentives such as a free cholesterol test or a voucher for a free lunch if you have trouble recruiting donors. NOTE: IRB approval may be required before drawing—check your local regulations.
5. SMILE can provide you with a sample health questionnaire. Typical blood bank donor screening questionnaires can also be used.
6. Some analytes have gender or age based partitions. For example, hematology RBC parameters have different ref ranges for men and women. You must collect 20 samples in each partition.

Transference with Verification

	If	Then
1	≥ 90% of samples are within the reference range	• The reference range is verified.
2	< 90% of samples are within the reference range	• Re-evaluate the range being verified. • Re-evaluate the healthy volunteer qualifications. • Collect and evaluate 20 additional samples.
3	≥ 90% of the additional samples are within the reference range	• The reference range is verified.
4	< 90% of the additional samples are within the reference range	• Proceed with step II below (Establishment of Reference Ranges)

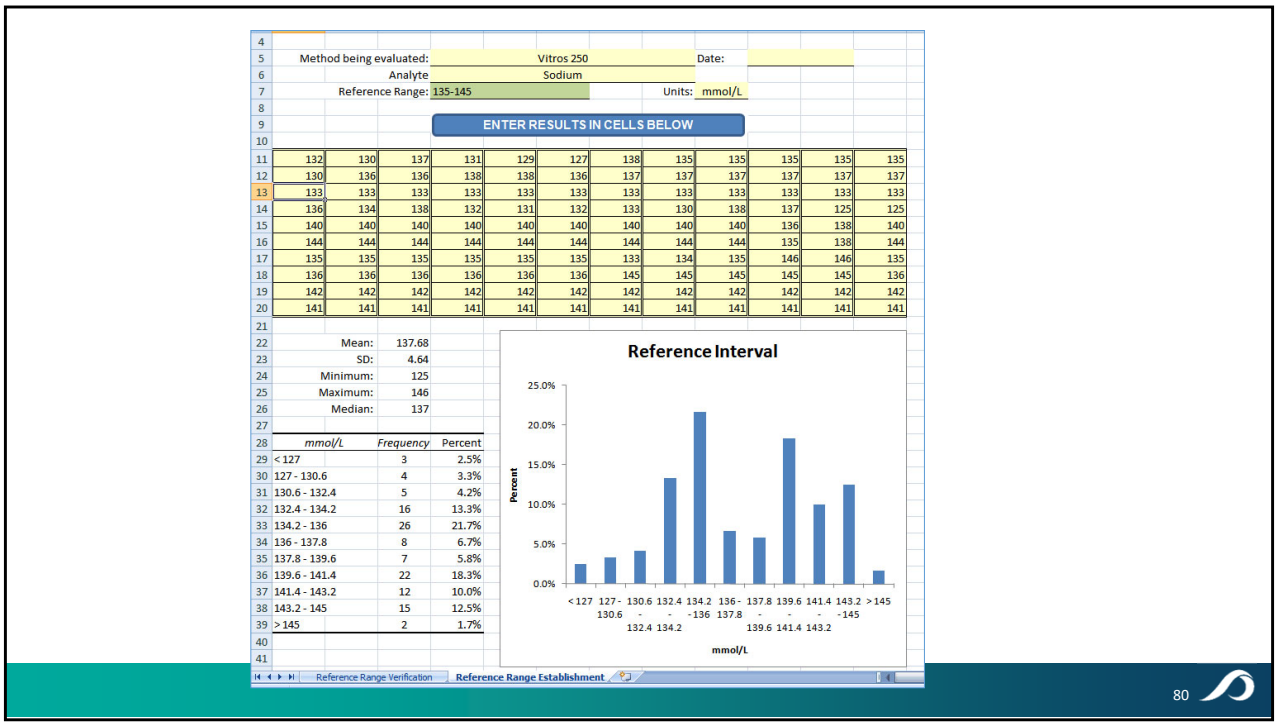
77



SMILE is a calculation tool that can be used to assess the transfer of reference values. Simply enter your proposed range and raw data. The tool will tell you if the results are acceptable or unacceptable. SMILE All tools are available in our website www.psmile.org.

Establishment of Reference Ranges

1. Qualify healthy volunteers. This can be done through a questionnaire or health assessment.
2. Obtain samples from 120 healthy participants for each range to be established.
3. Test each sample immediately after collection and evaluate.



SMILE has developed a spreadsheet tool that will do this assessment for you. Enter the raw data in the spreadsheet and conducted nonparametric assessment and give the reference range

Transference of Reference Ranges **without** verification

- CLSI guidelines permit the “transference” of established reference intervals without verification.

- Things to consider:
  - Similarity of geographics and demographics.
  - Similarity of test methodology.
  - Sound clinical judgment and consultation with local medical professionals.
  - Approval by the laboratory medical director is required and must be documented.

Important points to consider when using this approach

- The Medical Director is charged with the approval of reference ranges.
- Documentation is required and needs to include at least:
  - 1) source and reasons for range adoption
  - 2) written plan of review—including possible verification over time
- Usually only recommending for pediatric populations

82

The medical director, using clinical judgment, personal experience and research, would develop working reference ranges for all identified partitions considering each analyte separately.

Example Validation Summary

ABC Laboratory
Hospital Complex
123 Big Road
City-Township, Country

Validation Summary Report

Purpose: Validation

Description of Equipment/Process:
Equipment/Process: **Vitros 950 Chemistry Analyzer**
Serial Number: **999435**
Location: **ABC Lab, City-Township, Country**
Date: **18 June – 20 August 2008**
FDA Approval Status: Approved

Procedure:

Refer to the **ABC Lab** Validation Plan for **Vitros 950 Chemistry Analyzer**

Results: All raw data reports and statistical analysis can be found in the **Vitros 950 Chemistry** Validation binder.

1. Precision- refer to tab A

Analyte	Expected Results		Observed Results		Acceptability
			Between Day		
	Manufacturer's Precision	33% of CLIA	Normal Control CV%	Abn Control CV%	
ALT	3.3%	6.6%	3.8%	4.3%	Acceptable
AST	3.1%	6.6%	3.1%	2.1%	Acceptable
Albumin	1.5%	3.3%	2.8%	2.6%	Acceptable

Analyte	Expected Results		Observed Results		Acceptability
			Within Run		
	Manufacturer's Precision	25% of CLIA	Normal Control CV%	Abn Control CV%	
ALT	2.6%	5%	1.0%	0.9%	Acceptable
AST	2.4%	5%	1.7%	0.4%	Acceptable
Albumin	1.0%	2.5%	0.6%	0.8%	Acceptable

Organize & File:
- Plan
- Report
- Raw Data

83

One thing that we always emphasize is the importance of summarizing your validation when it is completed. You should summarize in a concise, easy to follow Summary Report. SMILE also has a template available for use to design your validation summary. Always file this summary, along with all of the raw data, together in a place that is easily retrievable. Laboratory audits—whether they are done by DAIDS, WFHI, CAP or other accrediting agencies, will always ask to see your validation data.

QUESTIONS?

References

1. CLSI. User Verification of Precision and Estimation of Bias: Approved Guideline-Third Edition. CLSI document EP15-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2014.
2. CLSI. Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline Second Edition. CLSI document EP05-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2014.
3. CLSI. Measurement Procedure Comparison and Bias Estimation Using Patient Samples; 3rd ed. CLSI guideline EP09c. Wayne, PA: Clinical and Laboratory Standards Institute; 2018.
4. CLSI. Preliminary Evaluation of Quantitative Medical Laboratory Measurement Procedures Implementation Guide-1st Edition. CLSI document EP10-ED3-IG. Wayne, PA: Clinical and Laboratory Standards Institute; 2022.
5. CLSI. Defining, Establishing and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline-Third Edition. CLSI document EP28-A3c. Wayne, PA: Clinical and Laboratory Standards Institute; 2010.
6. CLSI. Evaluation of Total Analytical Error for Quantitative Medical Laboratory Measurement Procedures. 2nd ed. CLSI guideline EP21. Wayne, PA: Clinical and Laboratory Standards Institute; 2016.
7. EP Evaluator Release 12.0, David G. Rhoads Associates Inc., datainnovations.com/ep-evaluator-resources.
8. Westgard, James O., Basic Method Validation: Training in Analytical Quality Management for Healthcare Laboratories, 4th edition, 2020 Madison, WI 53717.
9. CLSI. Evaluation of Qualitative, Binary Output Examination Performance. 3rd ed. CLSI guideline EP12. Clinical and Laboratory Standards Institute; 2023.

Acknowledgements

The presenter would like to thank the following:

NIH Division of AIDS -Daniella Livnat

Johns Hopkins University School of Medicine - pSMILE

Dr. Lori Sokoll - Principal Investigator

Mark Swartz- Project Manager