# Draft Guidance: Assay Migration Studies for *In Vitro* Diagnostic Devices

**Sally A. Hojvat, Ph.D.**
**Stefanie Akselrod, MS.**
**Marina V. Kondratovich, Ph.D.**

January 13, 2010
FDA/Industry
Roundtable

# Introduction

**"*Assay Migration Studies for In Vitro Diagnostic Devices*" draft CDRH-CBER document published Jan. 5th 2009**

- **General Update**
- **In Practice Observations**
- **Statistical Observations**

# General Update

- Internal CDRH/OIVD team formed to ensure review consistency and adherence to guidance
- Scientific reviewers, instrument/software reviewers and statisticians
- CDRH:  30 PMA Supplements received and reviewed
- Includes 8 platforms, all serological assays to date
- Multiple pre-IDEs reviewed, including NAAT systems
- Comments to guidance under review

# In Practice Observations

- Comparison of original study design based on older CLSI documents with newer guidelines
- Difficulty in preparing reproducibility panels for super-sensitive NAAT assays
- Handling of equivocal results – presentation of data
- Need to closely monitor studies

# Precision

- Comparing “apples to apples”
  - Original validation panel members:

    Unlike new validation panel members few with concentrations “close to the cutoff”
  - For ultra sensitive NAAT assays ,”high negative” members difficult to prepare

- Difference in variability ($SD_{New}/SD_{Old}$) between the two systems
  - Statistical vs. clinical significance

# Equivocal Results

- Evaluation of extent of differences in results near cutoff between platforms
- Expect some degree of “bounce”
- Checking for bias between the platforms
- Separate data analysis for these samples

# Comparison Panels

- Advantage of using more than one “old” system
  - Averages variability between instruments
- Allow sufficient volume for repeats due to un-anticipated problems
  - Operator errors (i.e. handling of samples, incorrect reagent lots, incorrect calibration discs)
  - Instrument failures

# Migration Review Issues

Example 1:

Discordant or repeat testing performed, when new instrument discordant with old instrument …*the testing was repeated on the new instrument two more times and a 2/3 result was used* …

Bias introduced since repeat analysis was not carried out for concordant results

# Migration Review Issues, cont'd

Example 2:

Reproducibility data analysis:

*… Between-day %CV analysis is missing from the new instrument analysis table but is presented for the original instrument…*

- Include same components of variance
- Any outliers or failed runs should be footnoted under reproducibility table in labeling

# Migration Review Issues, cont'd

Example 3:

…The on board reagent stability study was conducted at 4-10°C, using an external rocker mechanism designed to mimic the storage chamber of the instrument…

BUT

…the instrument reagent chamber temperature is maintained at 8-12°C...

# Statistical Observations

1) Qualitative test

- Test with two outcomes (neg., pos.);
- Test with three outcomes (neg., equiv., pos.);
- Ultrasensitive test

2) Comments about regression analyses

# **Basic Error Model** (General Concept)

Individual measurement of a given sample K

**Result = True Value**
**+ Mean-Bias**
**+ Random-Bias$_K$**
**+ Random Error**

Mean bias
(depends on the method)

Random Interferences
(samples from
different patients)

Precision
(different testing
conditions-
runs, days,
operators, …)

**New system measurements are exchangeable with Old system measurements.**

For a sample, measurement of the New System (Y) is as another measurement ($X_2$) by the Old System

# Qualitative Test with Two Outcomes

**Cutoff for qualitative test:**

- THRESHOLD for the OBSERVED result for a sample above which the result for a sample is reported as positive and below which the result is reported as negative;

Three scenarios:

1. Cutoff is based on clinical performance (Non-diseased and Diseased subjects have some amounts of analyte)

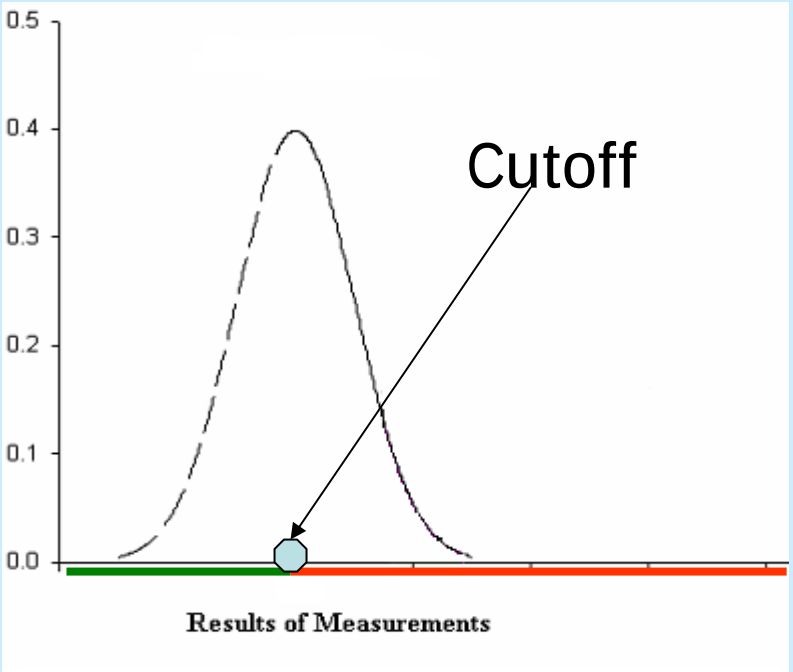
2.1 LoB is as cutoff (No analyte vs Analyte present), samples with zero concentrations produce positive signals

2.2 LoB=0 (ultrasensitive assay) samples with zero concentrations do not produce positive signals

# Qualitative Test with Two Outcomes

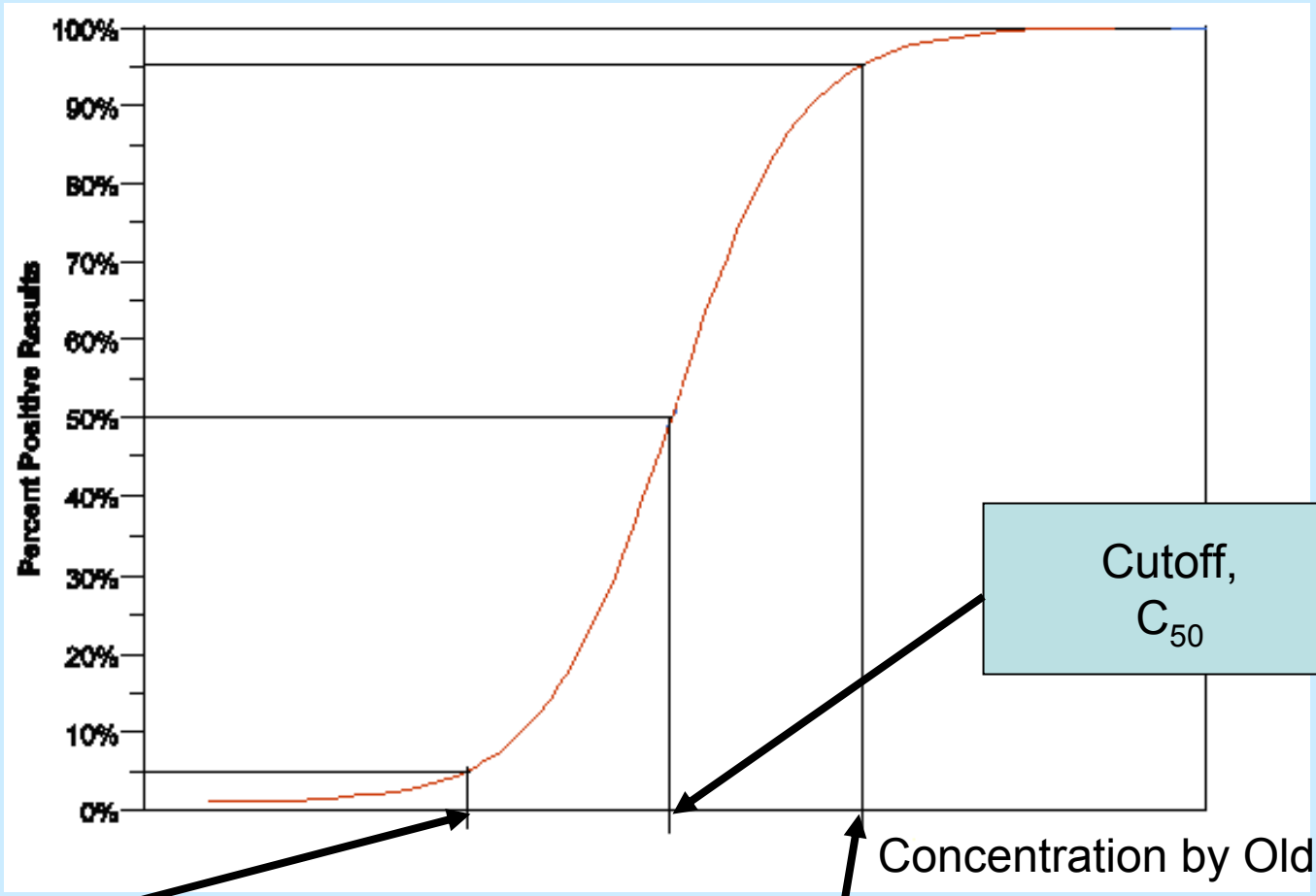
Cutoff is based on clinical performance
(Non-diseased and Diseased subjects have some amounts of analyte)

- Actual CONCENTRATION in a sample with this concentration is 50% positive and 50% negative ($C_{50}$) if a large series of repeated tests were performed



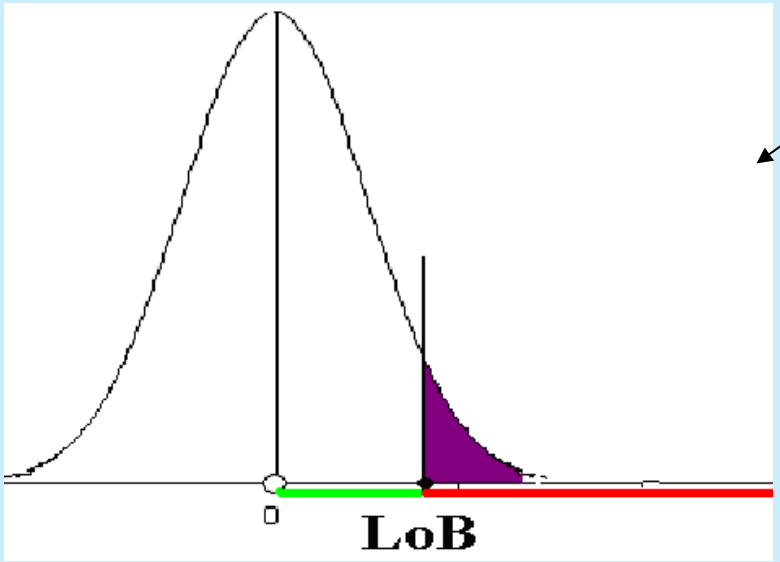

Assume that a distribution of measurement error is symmetrical.

# Qualitative Test



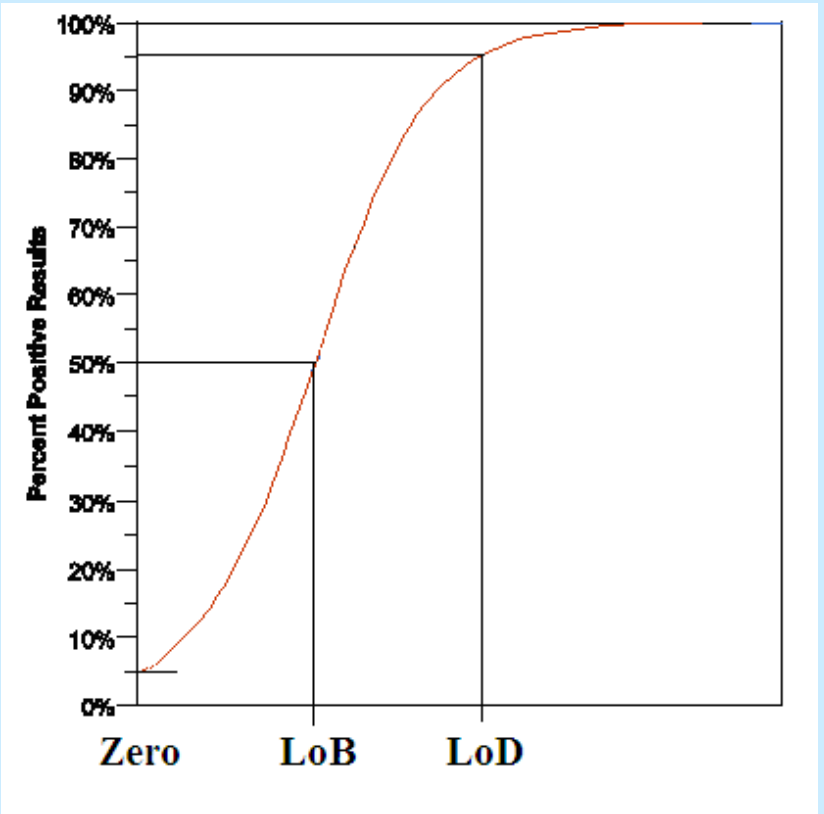
# Cutoff: Zero analyte vs Any amount of analyte

**Samples with zero concentrations produce positive signals**.




- Cutoff is based on the performance of the samples with zero concentration, Cutoff=LoB

- Percent of positive results for the samples with zero concentration is 5%;
- Percent of positive results for the samples with LoB concentration is 50%;
- Percent of positive results for the samples with LoD is 95%



Qualitative test

1) Precision around the cutoff
2) Systematic bias around the cutoff

# Types of Comparison Panel Samples

- Actual prospective patient samples
- Archived patient samples
- Banked patient samples
- Individual spiked or diluted patient samples
- Contrived matrix-specific samples

If some patient samples and concentrations may be rare

**Systematic bias in numerical outputs of Old vs New important around the cutoff.**

**Provide a scatter plot**

Samples with S/CO values around the cutoff:
Samples with S/CO values from approximately $C_5$ to approximately $C_{95}$

**Deming regression analysis is recommended**

- For each site separate
- For all sites combined

# Statistical Analysis for Comparison Panels

Positive and Negative Percent Agreements between Old and New Systems (PPA and NPA) (combining systematic bias and random error).

# Table of Agreement

| Site # | | Old System | | | |
|---|---|---|---|---|---|
| | | Negative | | Positive | |
| | | Strong and moderate negative | High negative, close to $C_5$ | Low positive, close to $C_{95}$ | Strong and moderate positive |
| New System | Negative | X | X | X | |
| | Positive | | X | X | X |

- For each site separate
- For all sites combined

(discordant results can only occur with samples close to the cutoff)

# Table of Agreement

| | | Old System | |
|---|---|---|---|
| | | **Negative** | **Positive** |
| **New System** | **Negative** | $A_1$ | $B_1$ |
| | **Positive** | $A_2$ | $B_2$ |
| **Total** | | $N_0$ | $N_1$ |

- Negative percent agreement: $A_1/N_0$ with 95% CI;
- Positive percent agreement: $B_2/N_1$ with 95% CI;

- For each site (95% CI by score method);
- For combined sites (95% CI by bootstrap)

# Qualitative Test with Equivocal Results

**Example:**

If S/CO < $E_1$ then "Negative",
If $E_1 \leq$ S/CO $\leq E_2$ then **"Equivocal",**
If S/CO > $E_2$ then "Positive"

A test with an equivocal zone of [$E_1$, $E_2$] can be considered as a test with **2 cutoffs**, the cutoff $E_1$ and the cutoff $E_2$.

**Systematic bias** in numerical outputs of Old vs New **around the cutoffs** is important.

**Scatter plot** of samples with S/CO values around the cutoff-

S/CO values from approximately $C_5$ of $E_1$ to approximately $C_{95}$ of $E_2$

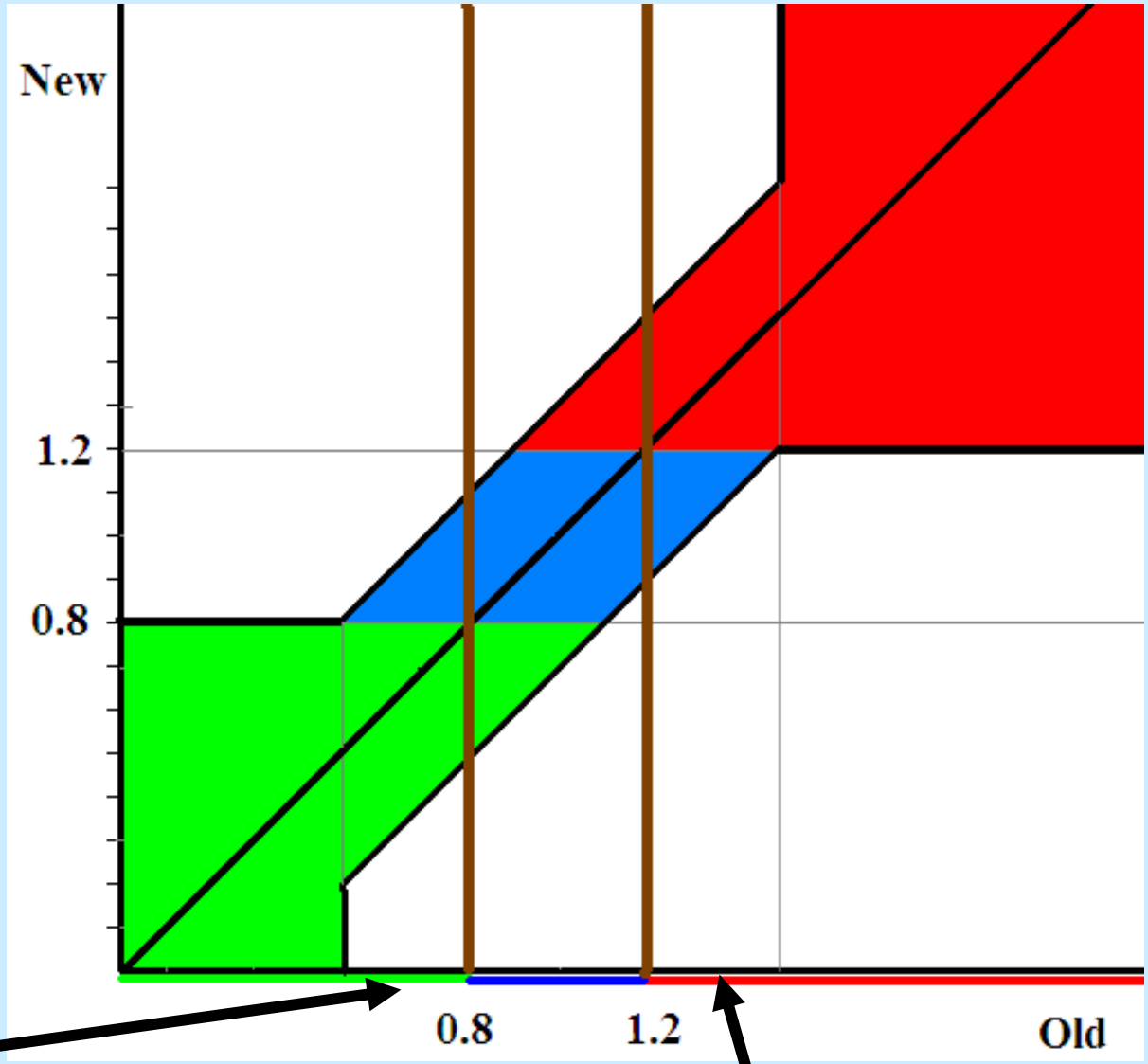
**Deming regression analysis**

For each site separate

For all sites combined

**Example:**
**$E_1$=0.8**
**$E_2$=1.2**
**(no re-testing),**

**In this example, SD is approximately constant around the 2 cutoffs (SD=0.1)**




Samples below cutoff $E_1$, close to $C_5$ of $E_1$

Samples above cutoff $E_2$, close to $C_{95}$ of $E_2$

# Table of Agreement

| Site # | | Old System | | | | |
|---|---|---|---|---|---|---|
| | | Negative | | Equivocal | Positive | |
| | | ≤ 0.50 | 0.51-0.79 | 0.80-1.20 | 1.21-1.49 | ≥ 1.50 |
| New System | ≤ 0.50 | X | X | | | |
| | 0.51-0.79 | X | X | X | | |
| | 0.80-1.20 | | X | X | X | |
| | 1.21-1.49 | | | X | X | X |
| | ≥ 1.50 | | | | X | X |

- For each site separate
- For all sites combined

# Table of Agreement

| | | Old System | | |
|---|---|---|---|---|
| | | **Negative** | **Equivocal** | **Positive** |
| **New System** | **Negative** | $A_1$ | $B_1$ | $C_1$ |
| | **Equivocal** | $A_2$ | $B_2$ | $C_2$ |
| | **Positive** | $A_3$ | $B_3$ | $C_3$ |
| **Total** | | $N_A$ | $N_B$ | $N_C$ |

- Negative percent agreement: $A_1/N_A$ with 95% CI;
- Positive percent agreement: $C_3/N_C$ with 95% CI;
- Among $N_B$ equivocal results by the Old system, there were ….

( Provide numbers (percents) of negative, equivocal and positive results by the New system).

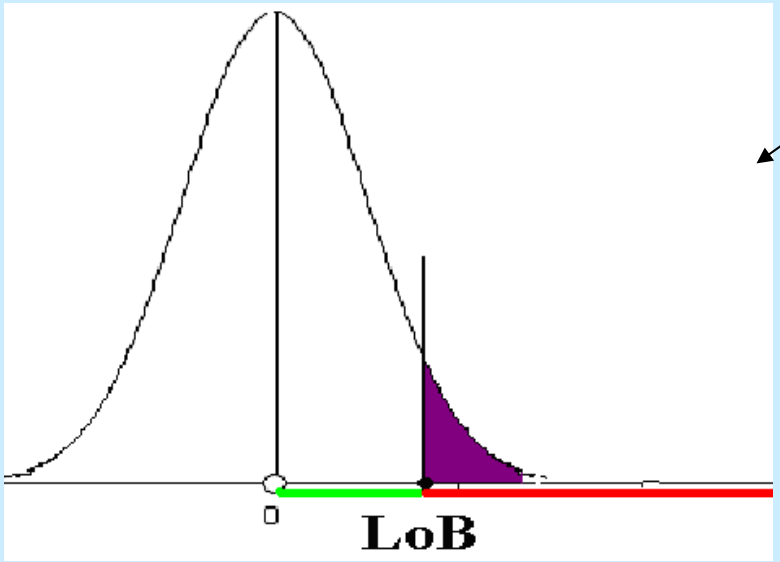
# Concentrations in the precision study for ultrasensitive assay

*Concentrations in the precision study (general case)*

At least three levels:

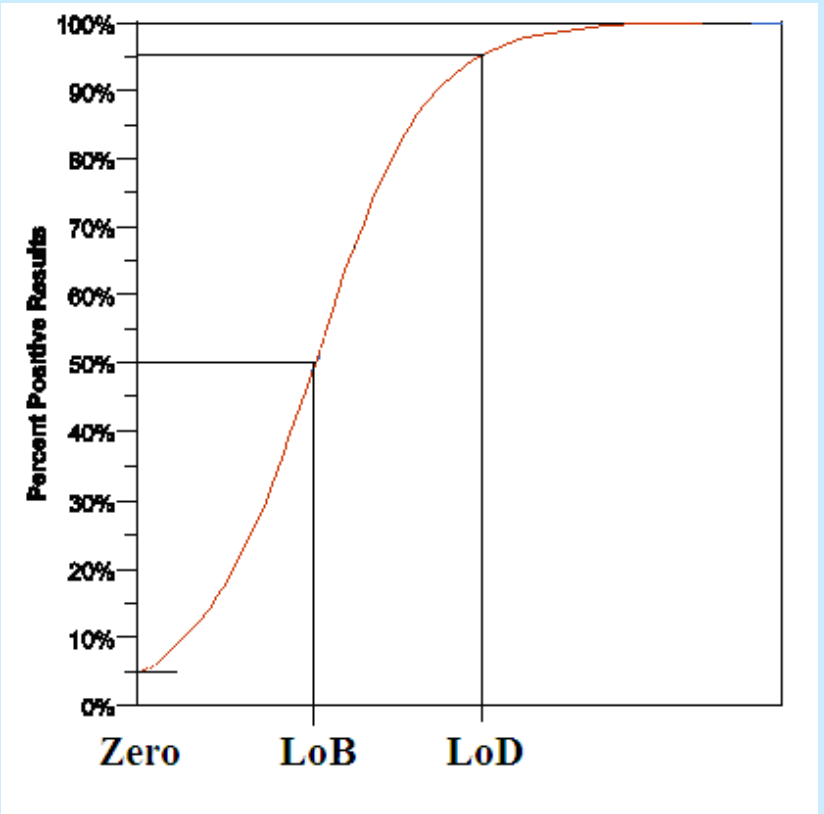
- High negative by Old System, (around $C_5$ of Old System);
- Low positive by Old System, (around $C_{95}$ of Old System);
- Moderate positive (positive results by Old System ≈ 100%)

# Most Common Case: samples with zero concentration produce positive signals.




- Cutoff is based on the performance of the samples with zero concentration, Cutoff=LoB

- Percent of positive results for the samples with zero concentration is 5%;
- Percent of positive results for the samples with LoB concentration is 50%;
- Percent of positive results for the samples with LoD is 95%



# Ultrasensitive Test: RT-PCR

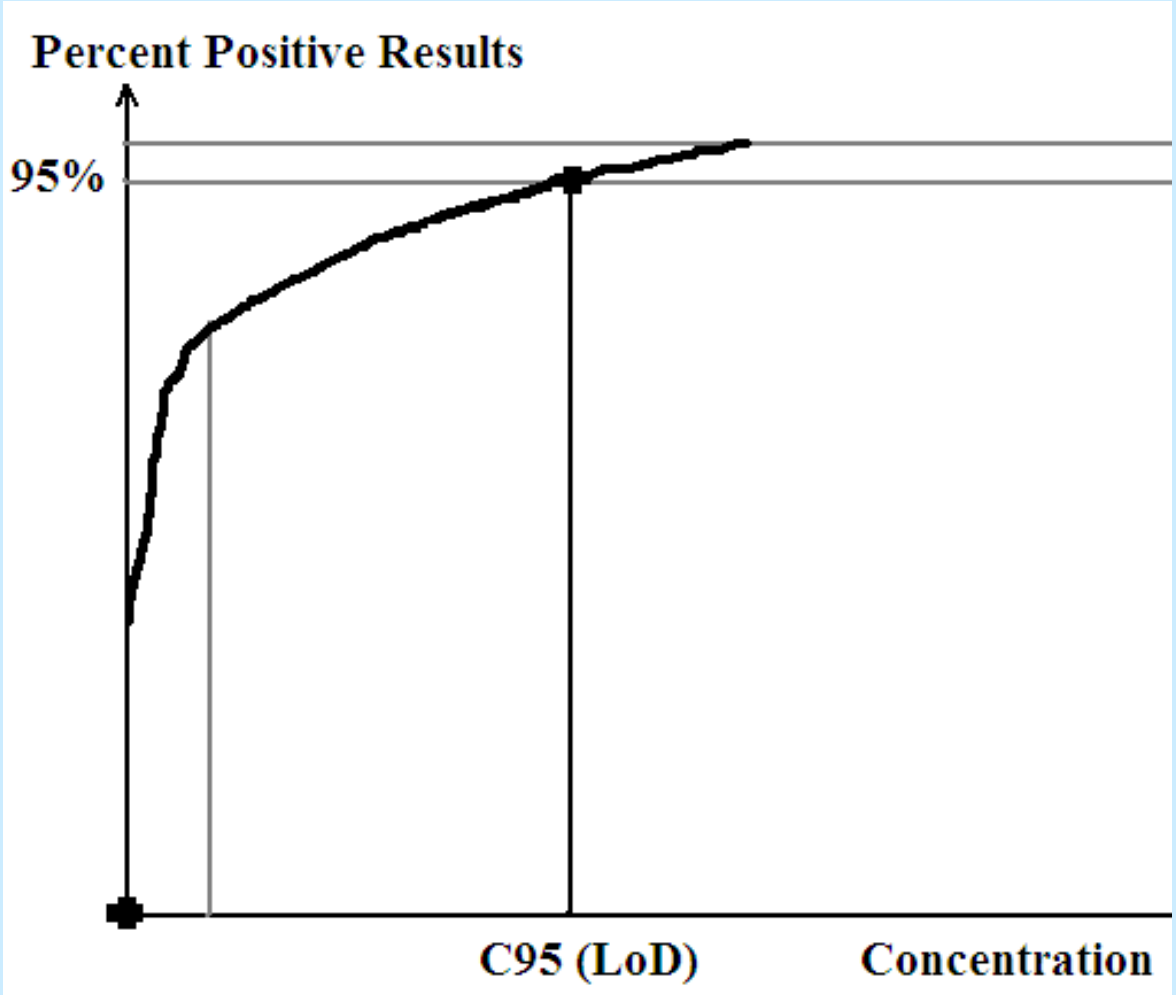
**Cutoff = 45 cycles**;

- Cutoff is not established based on the truly negative samples (zero concentration)

- If samples truly negative, all results are “Negative” => Type I error is close to zero

# Ultrasensitive Test

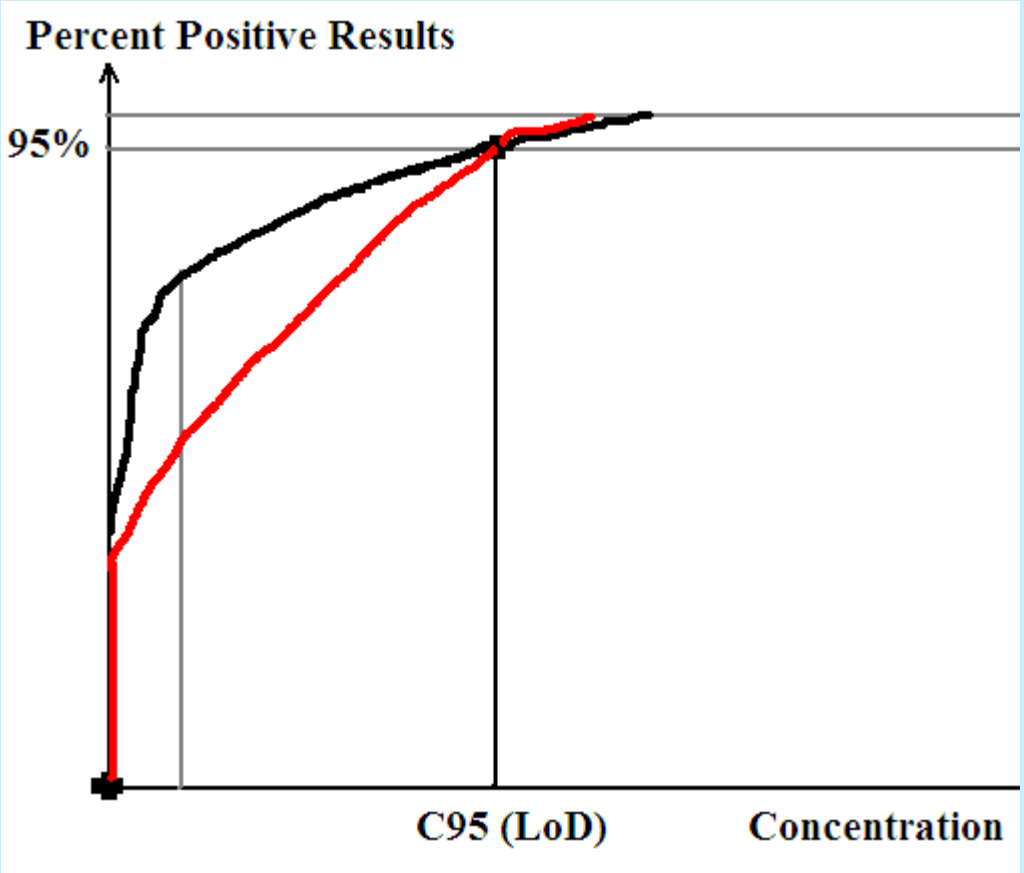
- Zero concentration has zero percent positive results;
- Concentration corresponding to Ct=45 is close to zero;
- $C_{95}$ (LoD)- concentration corresponding to Ct=38.



# Problem : $C_5$ not easy to prepare. Concentration is very close to 0 where large variability.

If only two points:

- **Zero concentration** percent of positive results is 0
- **LoD concentration** percent of positive results is 95%




then curves of % positive results of Old and New systems can be different (much uncertainty between two curves if only two points on these curves are similar)

# Modified Approach  Case A

If the percent of subjects from the intended use population with the test results less than $C_{95}$ (LoD) is less than **10%** of all subjects positive by the Old System **,** then no need for $C_5$.

Modified recommended concentrations for precision studies :

- truly negative sample
- $C_{95}$ (LoD) sample
- moderate positive

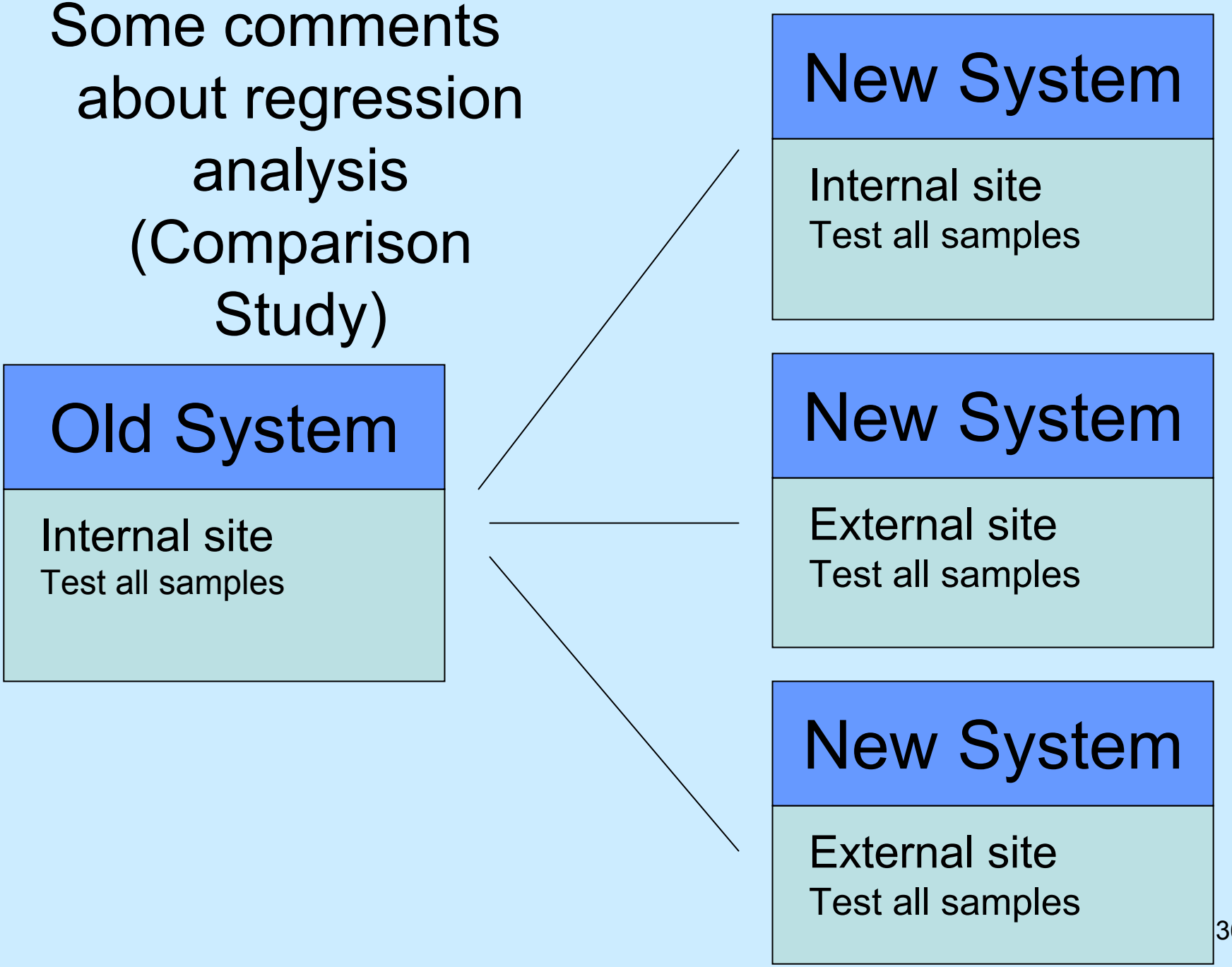
# Modified Approach Case B

If the percent of subjects from the intended use population with test results less than $C_{95}$ (LoD) is greater than or equal to **10%** of all subjects positive by the Old System, then a sample from the range $C_{20}$-$C_{80}$ should also be tested.

Modified recommended concentrations for precision studies :

- truly negative sample
- $C_{95}$ (LoD) sample
- Sample from range $C_{20}$-$C_{80}$
- Moderate positive

# Some comments about regression analysis (Comparison Study)



Four regression analyses:

1) X vs Y1 (site 1)
2) X vs Y2 (site 2)
3) X vs Y3 (site 3)
4) X vs (Y1+Y2+Y3)/3 (combined data)

Both methods have measurements errors =>

Passing-Bablok regression
or Deming regression

# Passing-Bablok

*Advantages*

1) Robust against outliers

2) Measurements errors of method X and method Y are the same type of distribution (not needed to be normal)

3) The variances of the measurement errors need not to be constant within the range but should remain proportional

$\sigma_Y/\sigma_X$ = constant

it can be that i) $\sigma_Y = Constant_1$, $\sigma_X = Constant_2$

ii) $\%CV_Y = Constant_1$, $\%CV_X = Constant_2$

*Disadvantages*

4) $\sigma_Y/\sigma_X = 1$ (not more than 1.5)

# Deming

*Advantages*

1) $\sigma_Y/\sigma_X = \lambda$ (not necessary 1)

2) If $\sigma_Y = \text{Constant}_1$, $\sigma_X = \text{Constant}_2$
then Deming regression (general)

if $\%CV_Y = \text{Constant}_1$, $\%CV_X = \text{Constant}_2$
then Weighted Deming

*Disadvantages*

3) Measurement errors should be normally distributed

Four regression analyses:

1) X vs Y1 (site 1)
2) X vs Y2 (site 2)
3) X vs Y3 (site 3)
4) X vs (Y1+Y2+Y3)/3
(combined data)

Passing- Bablok or
Deming (usually weighted)
$\sigma_Y/\sigma_X = 1$

$\sigma_Y/\sigma_X \neq 1$
Deming is recommended

Note:

- Old system (1 measurement at internal site)
- Old system measurements should be representative
- Follow EP9-A2 (different days of testing)

Systematic difference between different calibrations of the Old System???

- Advantage:
  - not only different days but different calibrations are considered
  - if two or more instruments of the Old system are considered;
  - see reproducibility of the Old system

# Next Steps

- ❑ Draft guidance – submit comments
- ❑ Revised draft to final guidance

**For questions contact**
**CDRH Sally Hojvat, Ph.D.,**
**phone: (301)-796-5455**
**e-mail: sally.hojvat@fda.hhs.gov**
**CDRH Marina Kondratovich, Ph.D.,**
**phone: (301)- 796-6036**
**e-mail: marina.kondratovich@fda.hhs.gov**