

Statistical methods in validation reports and QC

Jeremie Gras, M.D.
Medical Director, IPG
MB&C Course 2024



Institut de Pathologie et de Génétique
25, avenue Georges Lemaître B6041
Gosselies
Tél : +32 (0)71 47 30 47
accueil@ipg.be



Statistical methods in validation reports and QC

Plan

- 1) Regulatory context
- 2) Statistical concepts and methods
- 3) QC concepts & important metrics
- 4) Practical application to molecular diagnostics

Statistical methods in validation reports and QC

Institute for Pathology and Genetics (I.P.G.)

- ❖ Created 1958
- ❖ Located in the Brussels South BioPark in Gosselies (Biotech Hub)
- ❖ Our values:
 - Integrity – Equity – Respect – Communication – Professionalism
- ❖ Integrated diagnostic facility:
 - Pathology (>15 pathologists)
 - Center for Human Genetics (12 clinical geneticists)
 - Molecular Diagnostics (>20 PhDs)
- ❖ Revenues 2023: 50 Millions €
- ❖ ISO 15189:2012 accreditation BELAC 381-MED



Statistical methods in validation reports and QC

Plan

- 1) **Regulatory context**
- 2) Statistical concepts and methods
- 3) QC concepts & important metrics
- 4) Practical application to molecular diagnostics

1. Regulatory context

Regulation and norms regarding validation

- **ISO 15189: 2022**
- **IVD-R directive**
- **Practical directives for Pathology and Clinical biology**
- **Guidelines, reference articles**
- **Other regulations**

1. Regulatory context

ISO 15189:2022

- International norm
- ISO = International Organization for standardization
- Norm for Medical Laboratories, certainly non- specific
- Mandatory accreditation for most molecular tests in Belgium
- ISO 15189: 2022 mandatory starting Feb 2024 !

1. Regulatory context

ISO 15189:2022, validation in general

7.3 Examination processes

7.3.1 General

- a) The laboratory shall select and use examination methods which have been **validated** for their intended use to assure the clinical accuracy of the examination for patient testing.

NOTE Preferred methods are those specified in the instructions for use of in vitro diagnostic medical devices or those that have been published in established/authoritative textbooks, peer-reviewed texts, or journals, or in international and national consensus standards or guidelines, or national or regional regulations.

- b) The **performance specifications** for each examination method shall relate to the intended use of that examination and its impact on patient care.
- c) All procedures and supporting documentation, such as instructions, standards, manuals and reference data relevant to the laboratory activities, shall be kept up to date and be readily available to personnel (see [8.3](#)).
- d) Personnel shall follow established procedures and the identity of persons performing significant activities in examination processes be recorded, including POCT operators.
- e) Authorized personnel shall periodically evaluate the examination methods provided by the laboratory to ensure they are clinically appropriate for the requests received.

1. Regulatory context

ISO 15189:2022, verification

7.3.2 Verification of examination methods

- a) The laboratory shall have a procedure to **verify** that it can properly perform examination methods **before** introducing into use, by ensuring that the required performance, as specified by the manufacturer or method, can be achieved.
- b) The performance specifications for the examination method confirmed during the verification process shall be those relevant to the intended use of the examination results.
- c) The laboratory shall ensure the extent of the verification of examination methods is sufficient to ensure the validity of results pertinent to clinical decision making.
- d) Personnel with the appropriate authorization and competence shall review the verification results and record whether the results meet the specified requirements.
- e) If a method is revised by the issuing body, the laboratory shall repeat verification to the extent necessary.
- f) The **following records** of verification shall be retained:
 - 1) performance specifications to be achieved,
 - 2) results obtained, and
 - 3) a statement of whether the performance specifications were achieved and if not, action taken.

1. Regulatory context

ISO 15189:2022, validation

7.3.3 Validation of examination methods

- a) The laboratory shall validate examination methods derived from the following sources:
- 1) laboratory designed or developed methods;
 - 2) methods used outside their originally intended scope (i.e. outside of the manufacturer's instructions for use, or original validated measurement range; third party reagents used on instruments other than intended instruments and where no validation data are available);

1. Regulatory context

ISO 15189:2022, validation

- 3) validated methods subsequently modified.
- b) The validation shall be as extensive as is necessary and confirm, through the provision of objective evidence in the form of performance specifications, that the specific requirements for the intended use of the examination have been fulfilled. The laboratory shall ensure that the extent of validation of an examination method is sufficient to ensure the validity of results pertinent to clinical decision making.
- c) Personnel with the appropriate authorization and competence shall review the validation results and record whether the results meet the specified requirements.
- d) When changes are proposed to a validated examination method, the clinical impact shall be reviewed, and a decision made as to whether to implement the modified method.
- e) The following records of validation shall be retained:
 - 1) the validation procedure used;
 - 2) specific requirements for the intended use;
 - 3) determination of the performance specifications of the method;
 - 4) results obtained;
 - 5) a statement on the validity of the method, detailing its fitness for the intended use.

1. Regulatory context

What is not in ISO 15189: 2022

- A specific template for validation and verification
- Statistical methods that need to be used
- The number of patients samples (classical question !), control materials, cell lines, ... that need to be run in a validation or verification file

1. Regulatory context

IVD-R

- **New European regulation for In Vitro Diagnostics**
- **Competent agency: FAGG – AFMPS (additional inspections)**
- **IVD-R will make lab developed tests (LDT) more difficult to use**

1. Regulatory context

IVD-R – classification

Class D

**High public health risk,
high personal risk**

Examples

- HIV 1/2,
- Hepatitis C virus
- Hepatitis B virus
- HTLV I/II
- Blood grouping ABO, Rhesus (including RHW1), Kell, Kidd and Duffy systems
- CHAGAS
- Syphilis (used to screen blood donations)

Class C

**High personal risk,
moderate to low public health risk**

- Syphilis (diagnosis only)
- Neonatal screening for metabolic disorders e.g. PKU
- Rubella
- Cancer markers
- Genetic tests
- Companion diagnostics
- Blood glucose meters/strips
- Blood gas analysers
- Self tests

Class B

**Moderate to low personal risk,
low public health risk**

- Thyroid function
- Clinical chemistry
- Self-test devices listed as *not* Class C -> Pregnancy, Fertility, Cholesterol tests; and detection of glucose, erythrocytes, leucocytes and bacteria in urine

Class A

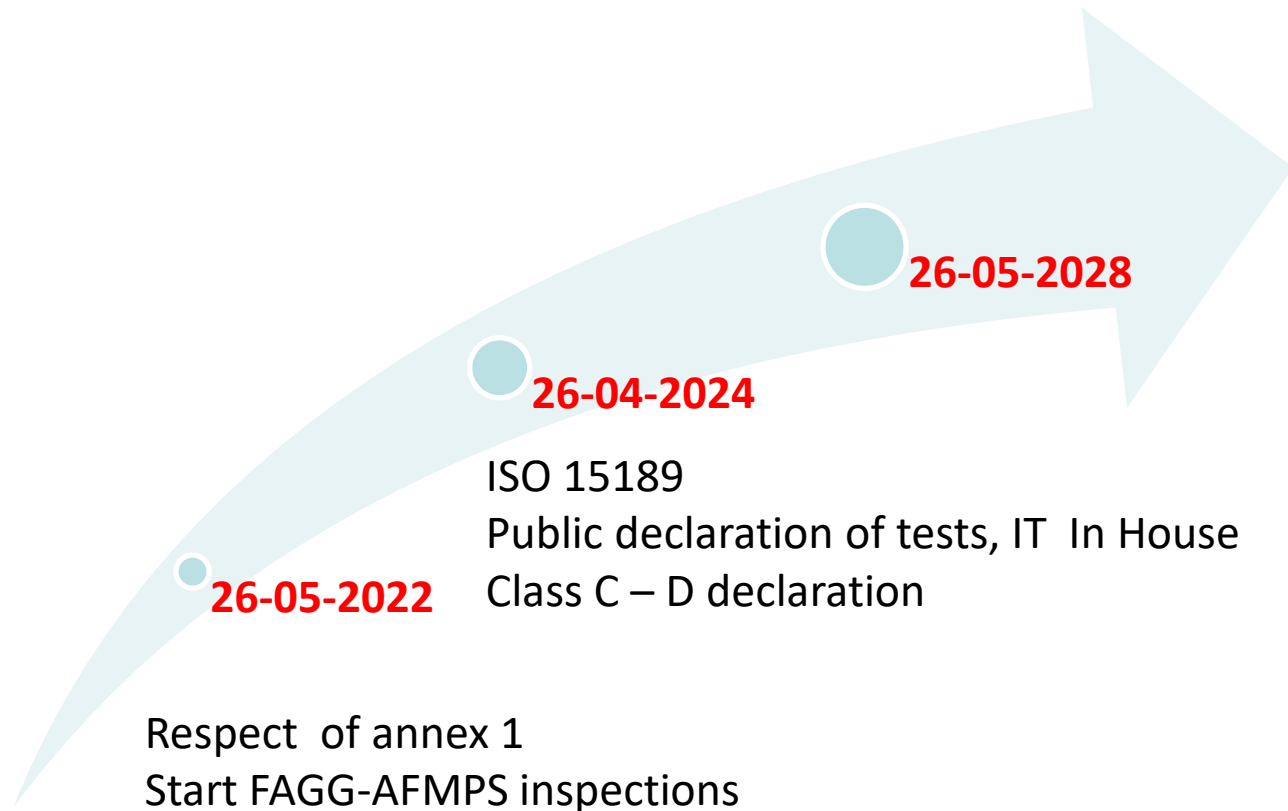
**Low personal risk, low
public health risk**

- Accessories
- Wash buffers
- Specimen receptacles
- Instruments
- Culture media

Source BSI

1. Regulatory context

IVD-R directive: timeline



1. Regulatory context

Guidelines and reference articles

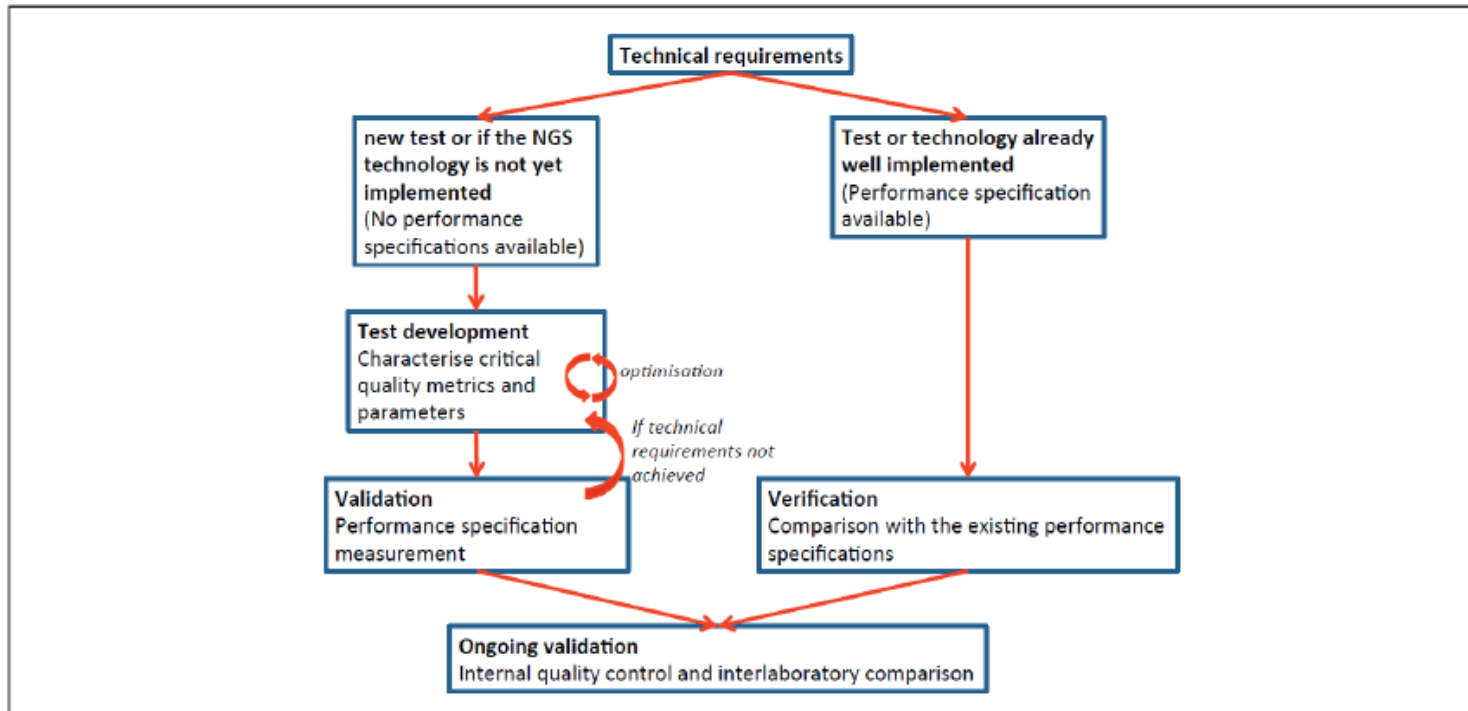


Figure 1 Workflow for the implementation of a NGS test

BELAC 2-405-NGS R3- 2021: guidelines for NGS validation

Statistical methods in validation reports and QC

Plan

- 1) Regulatory context
- 2) **Statistical concepts and methods**
- 3) QC concepts & important metrics
- 4) Practical application to molecular diagnostics

2. Statistical concepts and methods

What are statistics ? Why do we use statistics ?

- **Descriptive statistics** are used to characterize data
- **Statistical analysis** is used to distinguish between random and meaningful variations
- In the lab, we use statistics to **verify** method performance, **interpret** lab test results and **assure quality** of examinations procedures

2. Statistical concepts and methods

Needed for validation

- **Precision**
- **Accuracy**
- **Linearity – measuring range – limit of quantification**
- **Limit of detection (LOD) (analytical sensitivity)**
- **Analytical specificity**
- **Diagnostic performance (sensitivity, specificity, PPV, NPV, diagnostic efficiency)**

2. Statistical concepts and methods

Needed for verification

- **Precision**
- **Accuracy**
- **Verification of LoD and LoQ have been frequently requested by BELAC auditors**

2. Statistical concepts and methods

Table 2: Performance characteristics.

	FDA/CE-IVD Peer review multicentre publications	Home brew Adapted FDA/CE-IVD Adapted peer review multicentre publications
Precision (inter and intra run)	1 low positive sample, 1 high positive sample 3 replicates within 3 days. Preferentially from extraction.	1 low positive sample, 1 high positive sample 3 replicates within 7 days. Preferentially from extraction.
Accuracy	3 low positive samples, 3 high positive samples, 3 negative samples. If applicable, when selecting the positive samples, the genetic diversity should be taken into account. Preferentially from extraction.	10 low positive samples, 10 high positive samples, 20 negative samples. If applicable, when selecting the positive samples, the genetic diversity should be taken into account. Preferentially from extraction.
Linearity/ Measuring range/Limit of quantification	not necessary	Serial dilutions of min 5 log with 1 positive sample. 2 replicates in 2 runs. All log dilutions should be positive to be part of the measuring range.
Limit of detection/analytical sensitivity	not necessary	Can be concluded from linearity/measuring range experiment, followed by 20 measurements for lowest concentration with a confidence interval of 95 % (19/20 samples are positive)
Analytical specificity	not necessary	20 negative samples. If applicable: for microbiological tests, also analyse samples with microorganisms genetically related, unrelated but frequently detected in the same matrix, or presenting similar symptoms; for haematological tests, analyse samples from other haematological pathologies and healthy controls.

Definitions: Low positive sample= LOQ lowest concentration; High positive sample= LOQ highest concentration



2. Statistical concepts and methods

A few definitions:

Precision

- Closeness of agreement between independent results of measurements obtained under stipulated conditions

Repeatability

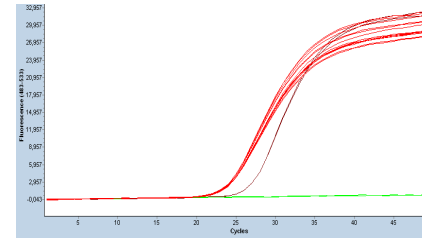
- Closeness of agreement between independent results of measurements obtained under the same operating conditions
 - Same operator
 - Same instrument
 - Same day
 - Same reagent lot
 - ...

2. Statistical concepts and methods

PRECISION: example- intra-run repeatability on 13 replicates

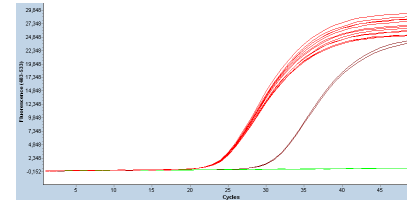
ABL

On Cq	1,13 %
On AQ	2,11 %
On AQ / ABL	N/A



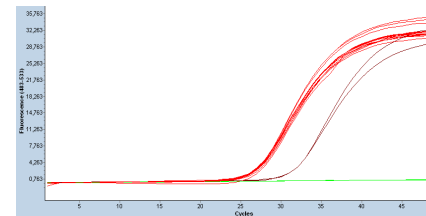
WT1

On Cq	0,21 %
On AQ	3,38 %
On AQ / ABL	4 %



BAALC

On Cq	0,37 %
On AQ	6,31 %
On QAQ/ ABL	6,48 %



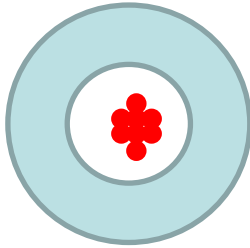
2. Statistical concepts and methods

A few definitions:

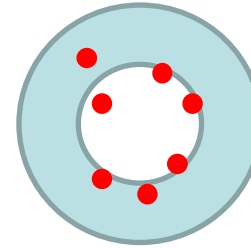
Reproducibility

- **Closeness of agreement between independent results of measurements obtained under changed conditions**
 - **Different operator**
 - **Different days**
 - **Different reagent lots**
 - **...**

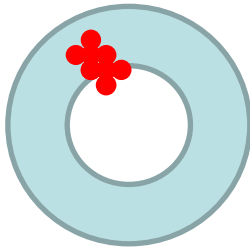
2. Statistical concepts and methods



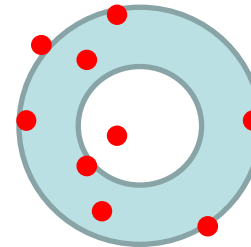
Accurate and precise
Low CV, low bias



Accurate, but not precise
High CV, low bias



Not accurate, but precise
Low CV, high bias



Not accurate and not precise
High CV, high bias

2. Statistical concepts and methods

Statistical concepts and methods

1. Type of data
2. Estimation of the centrality of data
3. Estimation of the variation of data
4. Confidence intervals
5. Regression
6. Probits
7. Bayesian statistics

2. Statistical concepts and methods

1. Type of data:

- **Quantitative data**
- **Categorical data**
 - **Nominal data**
 - **Ordinal data (scale data)**

2. Statistical concepts and methods

1. Type of data – applied to molecular tests

	Description	Examples
A	Quantitative tests. The result can have any value between two limits (including decimals).	Determination of methylation load (%); characterization of a mosaic mutation; heteroplasmy of mitochondrial variants.
B	Categorical tests where the quantitative signal is placed into an ordinal series to give the final result.	Sizing a PCR product; determination of triplet repeat size (FRAXA, Huntington disease, etc.)
C	Categorical tests where the quantitative signal is placed into one of a limited series of predefined categories to give the final result.	Determination of copy number using PCR or MLPA: exon deletion / duplication in <i>BRCA1</i> ; <i>PMP22</i> gene dosage in CMT and HNPP;
D	Qualitative tests where the true quantitative signal can have one of many possible values, but the required result can only have one of two possible values.	Mutation scanning for unknown mutations e.g. by sequencing or high resolution melt.
E	Qualitative [binary] tests where the true quantitative signal can only have one of two possible values	Genotyping for a specific mutation e.g. <i>CFTR</i> Phe508del in cystic fibrosis or <i>HFE</i> Cys282Tyr in hemochromatosis.

Mattocks CJ et al.
Eur J Hum Genet 2010; 18: 1276 – 1288.

2. Statistical concepts and methods

2. Estimation of centrality of data:

- **Mean(s)**
- **Mode**
- **Midrange**
- **Median**

2. Statistical concepts and methods

2. Estimation of the centrality of a dataset

- **Arithmetical Mean**

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

Arithmetical mean use *the sum* of the values

2. Statistical concepts and methods

2. Estimation of the centrality of a dataset

Geometrical mean

$$\bar{x}_g = \sqrt[N]{x_1 \cdot x_2 \cdot x_3 \cdots x_N} = \sqrt[N]{\prod_{i=1}^N x_i}$$

Geometrical mean use *the product* of the values

2. Statistical concepts and methods

2. Estimation of the centrality of a dataset

Harmonic mean

$$\bar{x}_h = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \dots + \frac{1}{x_N}} = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}}$$

Harmonic mean is appropriate to estimate the average *of rates*

2. Statistical concepts and methods

2. Estimation of the centrality of a dataset

Other means:

- Root square mean

$$\bar{x}_{rms} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_N^2}{N}} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$$

- Weighted mean

$$\bar{x}_w = \frac{\sum_{i=1}^N x_i \cdot w_i}{\sum_{i=1}^N w_i}$$

2. Statistical concepts and methods

2. Estimation of the centrality of a dataset- other than mean(s)

Mode:

- Mode is the value that occurs the most often

Midrange:

- Midrange is the mean of the highest and lowest values

2. Statistical concepts and methods

Median

The median is the value for which half of the remaining values are above and half are below it.

- *In an ordered array of 15 values, the 8th value is the median.*
- *If the array has 16 values, the median is the mean of the 8th and 9th values*

2. Statistical concepts and methods

2. Estimation of the centrality of a dataset

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G
1							
2			<u>Exemple 1</u>		<u>Exemple 2</u>		
3			1		22		
4			2		24		
5			3		26		
6			4		27		
7			5		31		
8			6		24		
9			7		25		
10			8		28		
11			9		29		
12			10		31		
13							
14							
15		Arithmetic mean	5,50		26,70		
16		Geometric mean	4,53		26,54		
17		Harmonic mean	3,41		26,39		
18		Median	5,50		26,50		
19							
20					=moye		
21							
22							
23							

The formula bar shows the formula `=moye` in cell E20. A dropdown menu is open, showing the following options:

- MOYENNE
- MOYENNE.GEOMETRIQUE
- MOYENNE.HARMONIQUE
- MOYENNE.RANG
- MOYENNE.REDUITE
- MOYENNE.SI
- MOYENNE.SI.ENS



2. Statistical concepts and methods

3. Estimation of variability

- Variance
- SD
- CV (RSD)

2. Statistical concepts and methods

3. Estimation of variability

- Two sets of data may have similar means, but otherwise be very dissimilar
- What is the **Variance** ?
 - “Variance is the mean of the squared differences between single data points and the mean of the array”
 - Variance has no units

2. Statistical concepts and methods

3. Estimation of variability

- Expression of variability= **Variance**

$$V = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

2. Statistical concepts and methods

3. Estimation of variability

- **Standard deviation (SD or sigma)**

$$\sigma = \sqrt{V} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$

2. Statistical concepts and methods

3. Estimation of variability

- **Standard deviation (SD or sigma)**
 - SD is the square root of the variance
 - SD has also no units

2. Statistical concepts and methods

3. Estimation of variability

- **Coefficient of variation (CV)**
 - **CV is the expression, in percentage, of the division of the standard deviation by the mean**
 - **Sometimes called the relative standard deviation (RSD)**

2. Statistical concepts and methods

3. Estimation of variability

- **Coefficient of variation (CV)**

$$CV = \frac{\sigma}{\bar{x}} \cdot 100$$

2. Statistical concepts and methods

4. Confidence intervals

- A confidence interval is a margin of error that indicates how precise an estimate is
- Generally, 95 % confidence intervals are used
- We can define 95 % CI for means, proportions, and other estimates

2. Statistical concepts and methods

4. Confidence intervals- for a mean

- To compute CI for a mean, we need first to compute the standard error (SE) of the mean
- SE of the mean = SD/\sqrt{n}
- As n increases, SE decreases and the precision is greater

2. Statistical concepts and methods

4. Confidence intervals- for a mean

- **95 % CI** for a mean:

from $- 1,96 \times SE \text{ (mean)}$ *to* $+ 1,96 \times SE \text{ (mean)}$

- **95 % CIs** are the most used
- For **90 % CI**, use **1,64** instead of **1,96**
- For **99 % CI**, use **2,58**

2. Statistical concepts and methods

5. Regression

- **Will be useful for:**
 - **Evaluation of linearity**
 - **Estimation of the limit of quantification**

2. Statistical concepts and methods

5. Regression

- Linear regression analysis is a function that will generate an equation for a straight line: $y = ax + b$

Where a is the slope of the line and b is the value of y when $x = 0$ (the *y-intercept*).

- The equation will give the expected value of y for each value of x

2. Statistical concepts and methods

Correlation versus linear regression

- Correlation using Pearson's correlation coefficient investigates *the strength* of a linear relationship between two continuous variables;
- It will give a correlation coefficient, noted r , and a p-value

2. Statistical concepts and methods

Correlation versus linear regression

- Linear regression investigates *the nature* of the linear relationship between two continuous variables;
- It will give an equation in the form of $y = ax + b$

2. Statistical concepts and methods

What is a P-value ?

- A P-value is a probability
- A P-value can take any value between 0 and 1
- P-value is the probability, given that the null hypothesis is true, of obtaining data as extreme or more extreme than the one that is observed
- Many statistical tests will give a P-value

2. Statistical concepts and methods

Null hypothesis

- **Null hypothesis: there is no difference between two comparison groups**

Interpretation of P-values

- **P values $< 0,05$ are generally considered significant;**
- **P values $> 0,05$ are generally considered non significant.**

2. Statistical concepts and methods

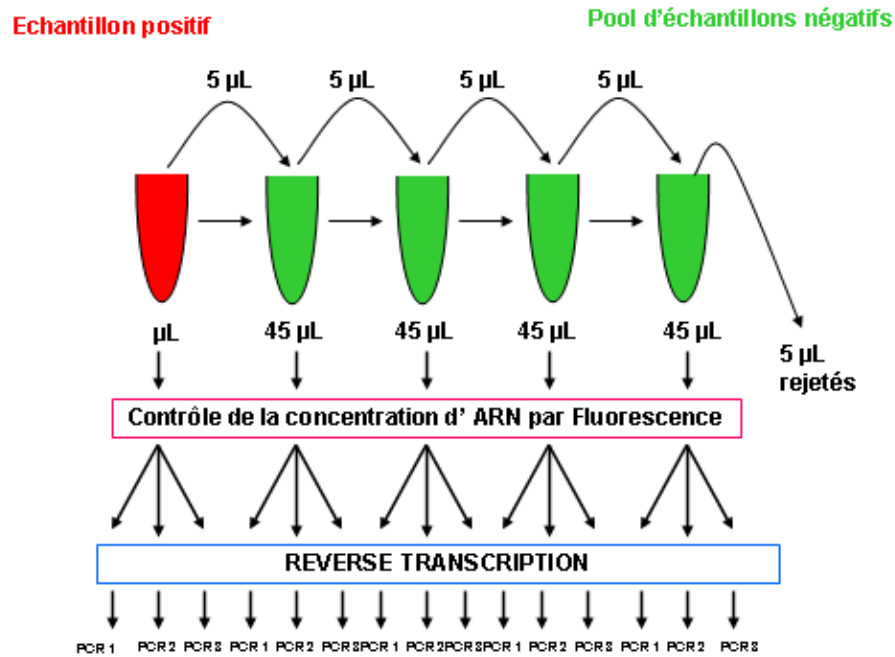
Other important general statistical concept one tailed and two-tailed tests

- = an option in many statistical tests
- *Two-tailed* or *two-sided* tests:
 - to be used when we cannot predict the direction of the difference between comparison groups
 - To be used in most of the cases
- *One sided* test:
 - to be used when we can predict the direction of the difference

2. Statistical concepts and methods

Linearity

- Example of linearity evaluation



2. Statistical concepts and methods

6. Probits

- **Probits = probability units**
- **Interesting statistical concept to estimate Limit Of Detection (LOD) at a little cost**

2. Statistical concepts and methods

LOD: methodology A

- Serial dilutions from a reference material (WHO or other)
- Estimation base on hit rate (proportion of positives)
 - 20 replicates located close to the estimated LOD
 - The dilution being tested 19 times positive on 20 is the LOD
 - Problem: Hard and costly

2. Statistical concepts and methods

LOD: methodology B

- Estimation based on Probits (Probability units)

2. Statistical concepts and methods

6. Probits

- Probits : “*Probability units*”
- Probit= Regression model
- Useful to estimate LOD when method gives a binomial answer
 - Binomial= detected or not detected
- Useful when the number of replicates is low
- Less costly

2. Statistical concepts and methods

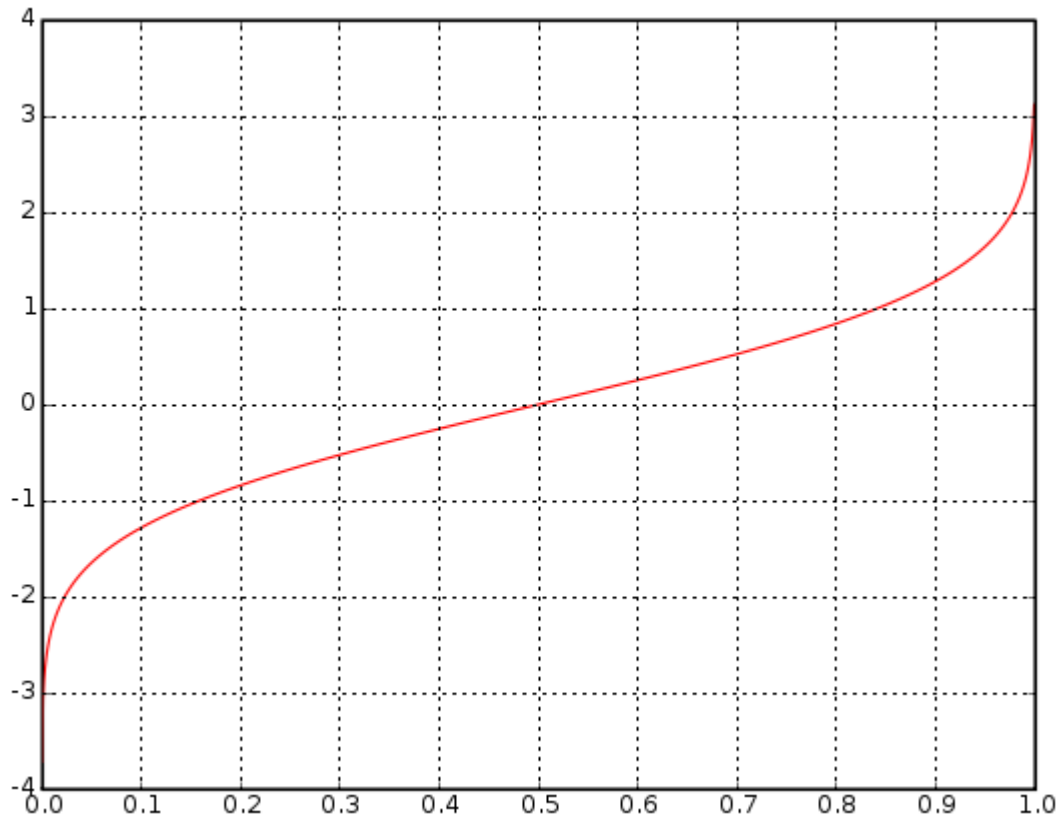
6. Probits

- Make serial dilutions
- Test these dilutions several times (from 3 to 8 to...)
- Calculate hit rate per dilution
- “Hit rate” = $H_i = N_{\text{pos}_i} / N_{\text{tot}_i}$
- Put the data on a graph: positivity on **Y** and concentration on **X**

2. Statistical concepts and methods

Probit approach uses IDF mathematical model :
Inverse Cumulative Distribution function

Probit

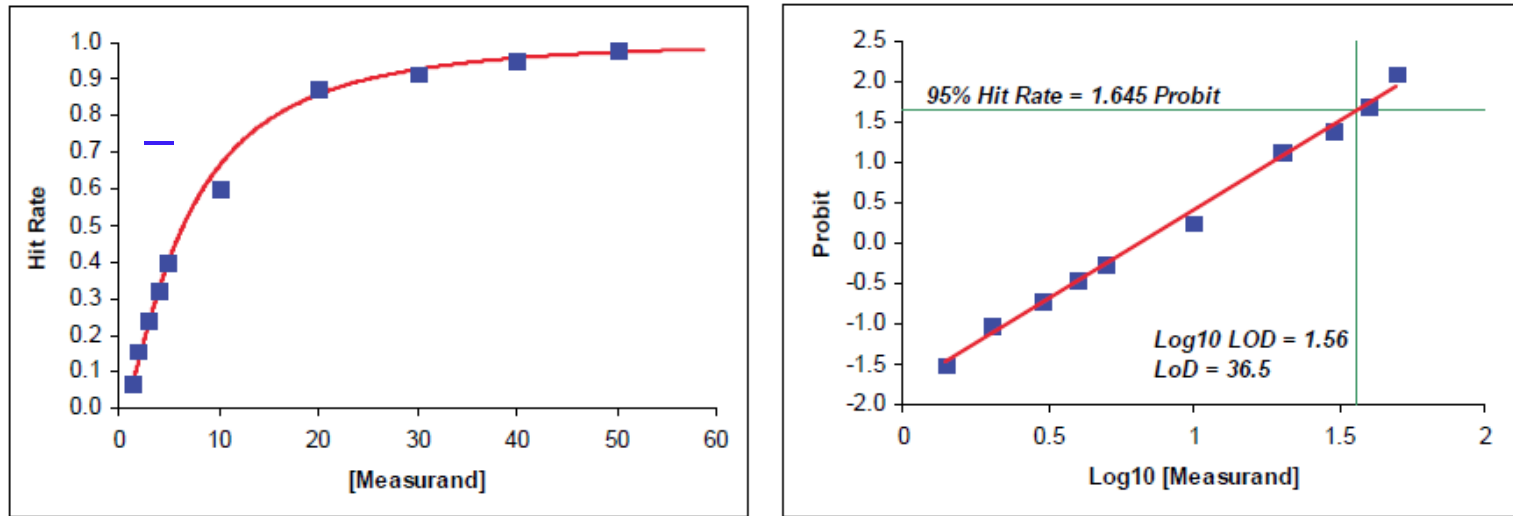


Proportion (hit rate)



2. Statistical concepts and methods

LOD/ Probit



Abbreviation: LoD, limit of detection.

Figure 3. Probit Analysis. These plots illustrate hypothetical experimental results (left hand plot) and regression analysis to determine the LoD for a molecular measurement procedure by the probit approach (right hand plot).

CLSI EP17-A2 reference document



2. Statistical concepts and methods

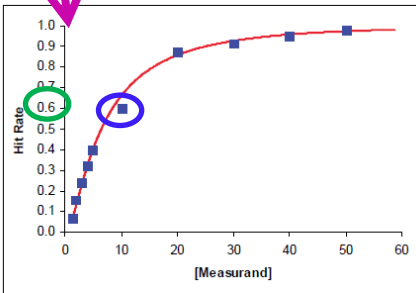
LOD/Probit

- **Probit analysis transforms the concentration/response curve into a straight line**
- **Probit= computed with statistical softwares (XLSTAT,...)**

2. Statistical concepts and methods

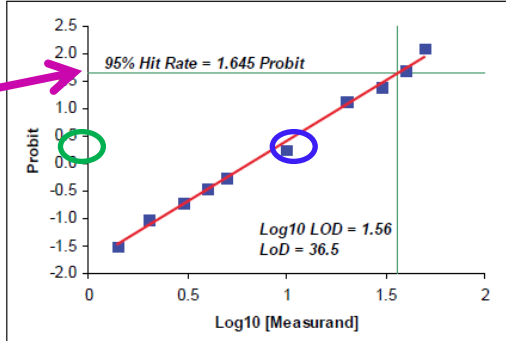
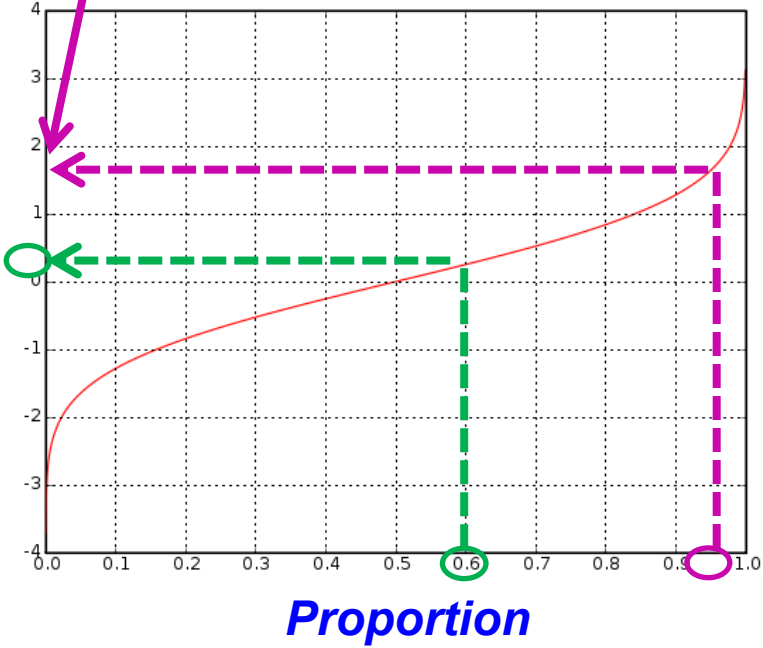
Probit approach

95 %
hit rate



Abbreviation: LoD, limit of detection.
Figure 3. Probit Analysis. These plots illustrate 1 regression analysis to determine the LoD for a mole (right hand plot).

1,645
Probit



Probit

Proportion



2. Statistical concepts and methods

LOD/Probit examples

Three HBV Viral Load Assay Formats

TABLE 1 Limits of detection^a

IU/ml	CAP/CTM			HP			TNAI		
	No. detected	No. not detected	% Detected	No. detected	No. not detected	% Detected	No. detected	No. not detected	% Detected
0	0	14	0	0	14	0	0	14	0
2.5	8	6	57	10	3	77	11	3	79
5	10	4	71	13	1	93	14	0	100
10	13	1	93	14	0	100	14	0	100
15	14	0	100	14	0	100	14	0	100
20	14	0	100	14	0	100	14	0	100
25	14	0	100	13	0	100	14	0	100

^a Probit results (95% confidence interval) were as follows: CAP/CTM, 10.2 IU/ml (6.8 to 27.8); HP and TNAI, not available.

The limits of detection of the assays were evaluated using dilutions of the WHO standard material as shown in [Table 1](#). Probit analysis predicts a limit of detection of 10.2 IU/ml (95% confidence interval = 6.8 to 27.8 IU/ml) for CAP/CTM. The same concentrations of diluted standard had inadequate resolution for probit analysis for HP and TNAI. However, 100% of the samples were detected at ≥ 10 IU/ml in the HP assay, and 100% of the samples were detected at ≥ 5 IU/ml in the TNAI assay.



2. Statistical concepts and methods

7. Bayesian statistics

- **Evaluation of the diagnostic accuracy**
- **Based on the exploitation of a contingency table**

2. Statistical concepts and methods

Evaluation of the clinical performance of a test

- The *sensitivity* of a test indicates the likelihood that the test will be positive when disease is present
- The *specificity* of a test indicates the likelihood that the test will be negative when disease is absent
- The *predictive value* of a test indicates the probability that the test result correctly classifies a patient

2. Statistical concepts and methods

Diagnostic sensitivity and specificity= Bayesian statistics

	Gold Standard		
	Disease present	Disease absent	
Positive test	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>	TP + FP
Negative test	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>	FN + TN
	TP + FN	FP + TN	



2. Statistical concepts and methods

Bayesian statistics: sensitivity

	Gold Standard		
	Disease present	Disease absent	
Positive test	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>	TP + FP
Negative test	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>	FN + TN
	TP + FN	FP + TN	

Sensitivity= sick patients who show a positive test

Sensitivity= $TP / (TP+FN)$



2. Statistical concepts and methods

Bayesian statistics: Specificity

	Gold Standard		
	Disease present	Disease absent	
Positive test	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>	TP + FP
Negative test	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>	FN + TN
	TP + FN	FP + TN	

Specificity = healthy people that show a negative test

$$\text{Specificity} = \frac{TN}{(FP+TN)}$$



2. Statistical concepts and methods

Bayesian statistics: Positive Predictive Value (PPV)

	Gold Standard		
	Disease present	Disease absent	
Positive test	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>	TP + FP
Negative test	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>	FN + TN
	TP + FN	FP + TN	

PPV= If the test is positive, what is the chance for the patient to have the disease ?

$$\text{PPV} = \text{TP} / \text{TP} + \text{FP}$$



2. Statistical concepts and methods

Bayesian statistics: Negative Predictive Value (NPV)

	Gold Standard		
	Disease present	Disease absent	
Positive test	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>	TP + FP
Negative test	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>	FN + TN
	TP + FN	FP + TN	

NPV= If the test is negative, what is the chance for the patient to really not have the disease?

$$NPV = TN / FN + TN$$



2. Statistical concepts and methods

Bayesian statistics: Diagnostic accuracy

Diagnostic accuracy = the percentage of patients that are correctly categorized by the test

$$DA = (TP+TN) / (TP+FP+FN+TN)$$

2. Statistical concepts and methods

Bayesian statistics example: NGS validation

Table 1. Comparison of NGS results with conventional methods.

Gene	Conventional Methods Result		Parameter	Agreement
	Mutation/Fusion Detected	Not Detected		
<i>EGFR</i> -NGS Result				
Mutation Detected	13	0	PPA	92.9%
Not Detected	1	86	NPA	100%
			OA	99.0%
<i>KRAS</i> -NGS Result				
Mutation Detected	74	5	PPA	97.4%
Not Detected	2	50	NPA	90.9%
			OA	94.7%
<i>NRAS</i> -NGS Result				
Mutation Detected	3	0	PPA	100%
Not Detected	0	54	NPA	100%
			OA	100%
<i>BRAF</i> -NGS Result				
Mutation Detected	9	2	PPA	100%
Not Detected	0	45	NPA	95.8%
			OA	96.4%
<i>ALK</i> Fusions-NGS Result				
Fusion Detected	4	1	PPA	100%
Not Detected	0	95	NPA	99.0%
			OA	99.0%
<i>ROS1</i> Fusions-NGS Result				
Fusion Detected	1	0	PPA	100%
Not Detected	0	99	NPA	100%
			OA	100%

NGS—Next Generation Sequencing; PPA—Positive percent agreement; NPA—Negative percent agreement; OA—Overall agreement.

2. Statistical concepts and methods

Statistical packages

- **Statistical packages are softwares that allow to compute statistical tests in a very fast way**
- **These include SAS, SPSS and Stata; they may be expensive, but there are student licences (for students)**
- **There are also add-ons for Microsoft Excel, such as Xlstat and Analyze – IT**

Statistical methods in validation reports and QC

Plan

- 1) Regulatory context
- 2) Statistical concepts and methods
- 3) **QC concepts & important metrics**
- 4) Practical application to molecular diagnostics

3. QC concepts and important metrics

QC concepts

- **Statistical Quality Control (QC) has long been used in clinical laboratories (clinical chemistry, immunoassays, ...)**
- **Theory is dating from 1970's- 1980's (seminal papers from JO Westgard)**
- **A “lifting” was applied to these concepts in 2005 with the publication of JO Westgard book “Six Sigma- quality control and design”**

3. QC concepts and important metrics

QC concepts

- In 2022, statistical QC (statistical treatment of QC data) is not a common practice in molecular diagnostics
- When applied, molecular statistical QC is sometimes limited to a Levey-Jennings graph with $\pm 3 Cq$ limits !

3. QC concepts and important metrics

QC concepts: P_{ed} and P_{fr}

▪ P_{ed} = probability of a QC rule to detect significant errors

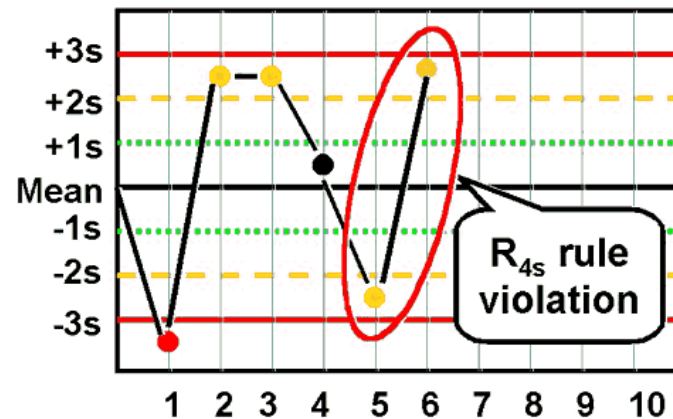
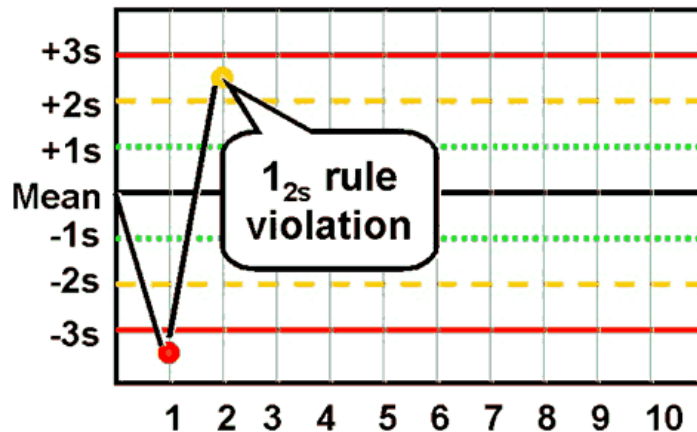
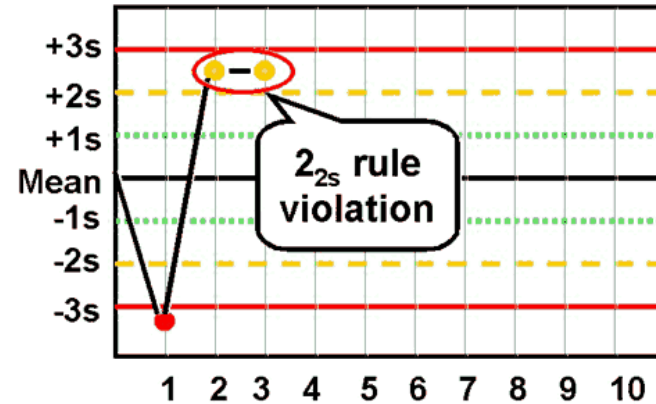
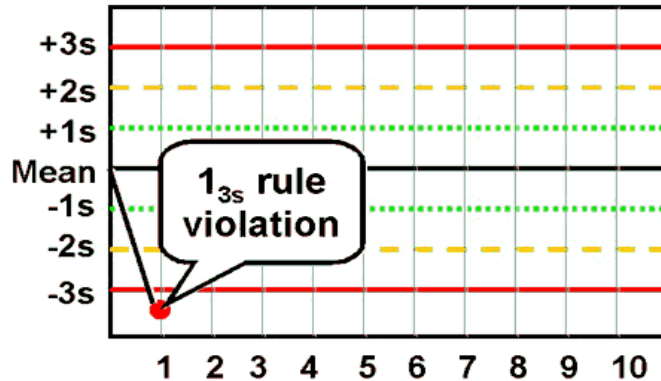
Goal for a statistical QC procedure: $P_{ed} > 0,9$ or $> 90\%$

▪ P_{fr} = probability of falsely rejecting results when there is no significant out-of-control condition

Goal for a statistical QC procedure: $P_{fr} < 0,05$ or $< 5\%$

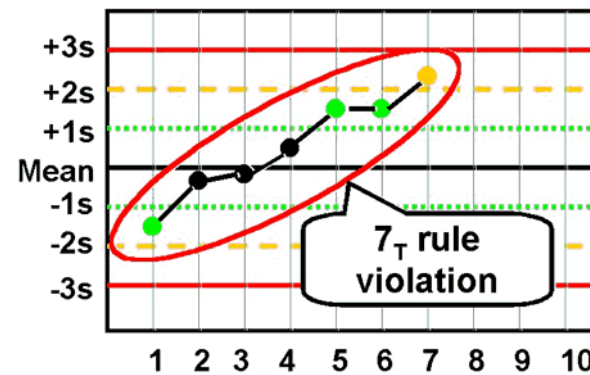
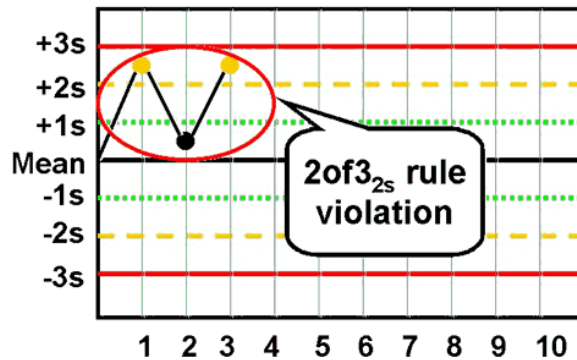
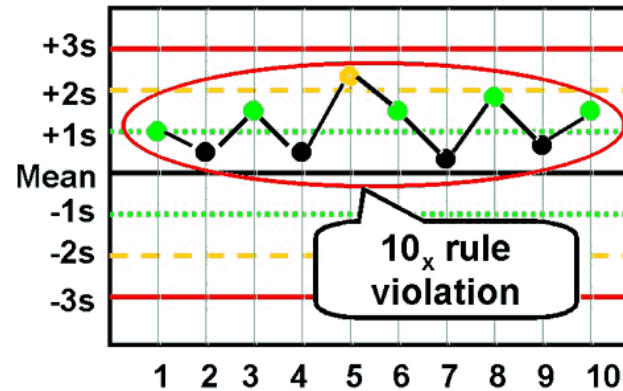
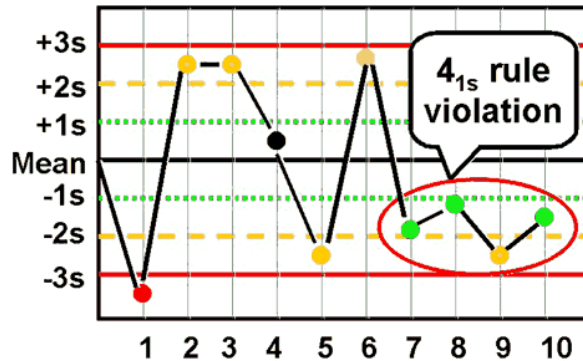
3. QC concepts and important metrics

QC rules (Westgard rules)



3. QC concepts and important metrics

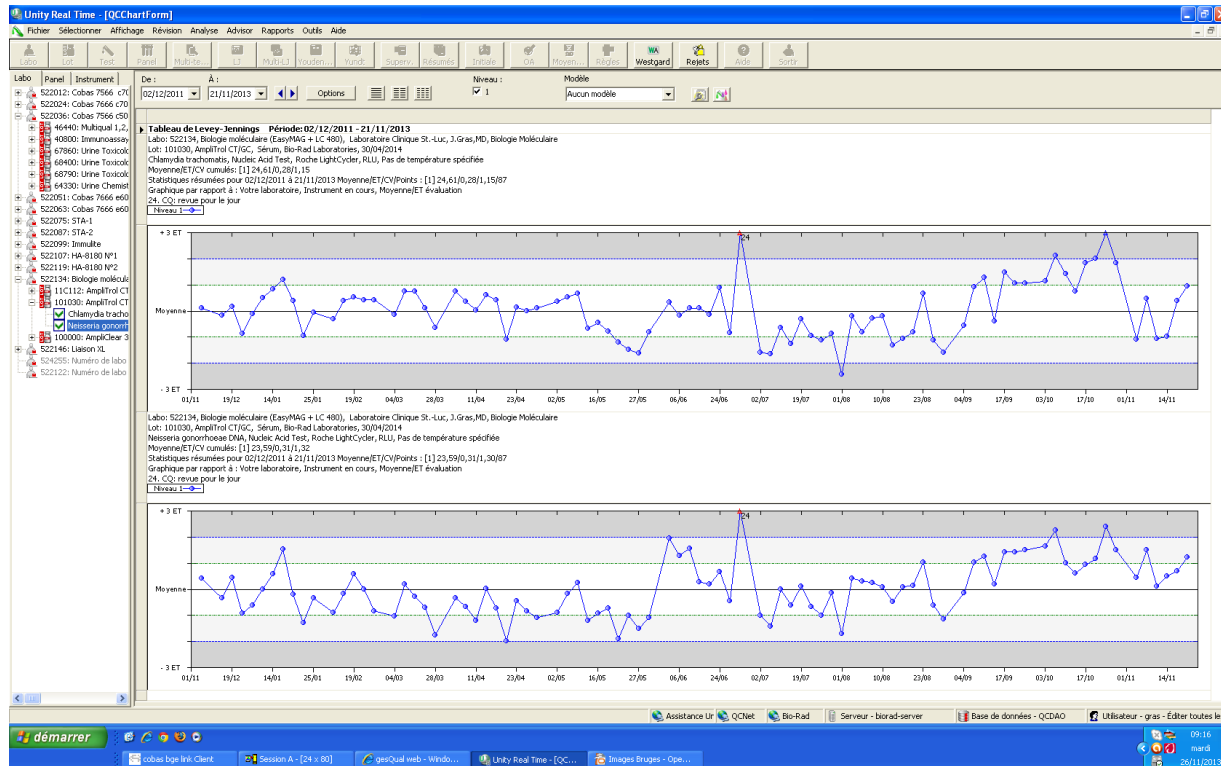
QC rules (Westgard rules)



3. QC concepts and important metrics

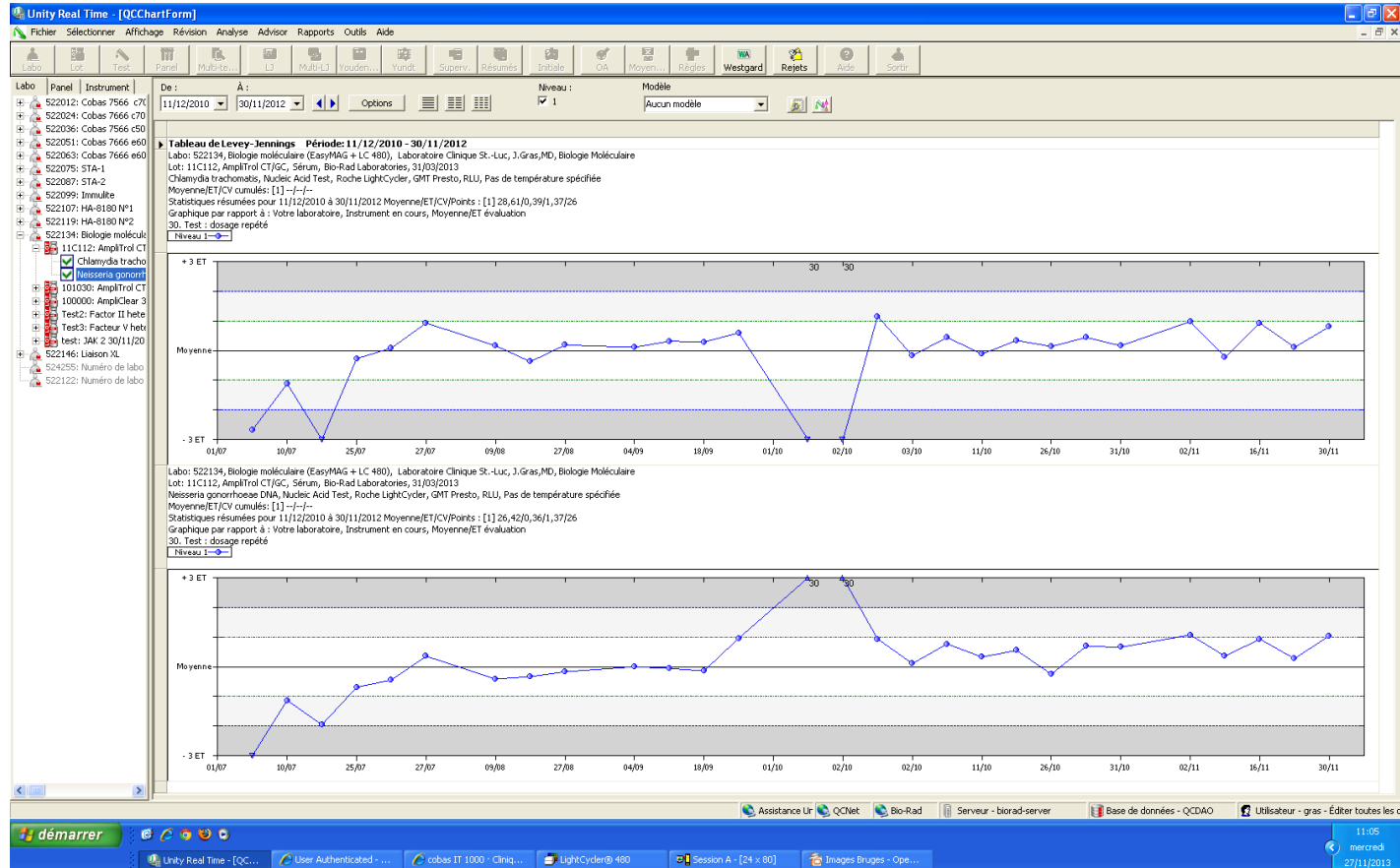
QC charts (Levey Jennings charts)

Plot of Cq value of positive control (CT and NG)



3. QC concepts and important metrics

QC charts (Levey Jennings charts) Violation of QC rules (CT/NG)



3. QC concepts and important metrics

- Traditional QC principles *are applicable* to molecular diagnostics:
- Minimal condition is to have a **continuous numerical value**:
 - Cq value
 - Copies number
- An **independent QC** (from DNA extraction) is preferable

Statistical methods in validation reports and QC

Plan

- 1) Regulatory context
- 2) Statistical concepts and methods
- 3) QC concepts & important metrics
- 4) **Practical application to molecular diagnostics**

4. Practical application to molecular diagnostics

qPCR: example of a verification file


	Qualité des résultats	IPG-VAL-347
	SARS-CoV-2- Perkin Elmer	Version n°: 001

Table des matières

I. GENERALITES	3
1. Description et champ d'application	3
2. Origine de la technique	3
3. Matrice et support	4
4. Méthodologie	4
5. Détails opérationnels	5
6. Contrôles internes pour la validation	5
7. Critères d'acceptabilité et interprétation des résultats	6
8. Limite(s) de la technique / Interférences	7
9. Directives en vigueur	7
10. contrôle de qualité interne (IQC)	7
11. Contrôle de qualité externe (EQC)	7
12. Divers	7
II. VALIDATION DU TEST	8
1. Répétabilité (+ calcul d'incertitude de mesure si justifié)	8
2. Fidélité intermédiaire (+ calcul d'incertitude de mesure si justifié)	8
3. Justesse / exactitude (+ calcul d'incertitude de mesure si justifié)	9
4. Sensibilité (+ calcul d'incertitude de mesure si justifié)	12
5. Limite de détection/ de quantification, seuil de détection (+ calcul d'incertitude de mesure si justifié)	12
6. Linéarité (+ calcul d'incertitude de mesure si justifié)	13
Conclusion. L'objectif concernant le coefficient de détermination est atteint car il est supérieur à 0.95.	15
7. Spécificité	15
8. Robustesse (+ calcul d'incertitude de mesure si justifié)	15
9. Analyse rétrospective	15
10. Analyse prospective	15
11. Autre	16
1. Vérification initiale du kit CE-IVD SARS-CoV-2 RT-qPCR reagent kit (Perkin Elmer)	16
2. Validation du flux des résultats du LC480 au CR dans DIAMIC	17
III. CONCLUSIONS	21
IV. MODIFICATION(S) APPORTEE(S) au DOSSIER de VALIDATION.	22
Titre :	22

4. Practical application to molecular diagnostics

qPCR verification: repeatability

1. Répétabilité (+ calcul d'incertitude de mesure si justifié)

Fidélité mesurée avec des résultats d'essai indépendants obtenus par la même méthode, sur les mêmes échantillons, dans le même laboratoire, avec le même opérateur et les mêmes équipements utilisés en même temps ou durant une courte période de temps.

Objectif. Concordance des résultats intra-run.

Méthodologie et échantillons. 2 échantillons positifs pour SARS-CoV-2 et un échantillon négatif pour SARS-CoV-2 ont été analysés en triplicat au cours d'un même run d'extraction sur Chemagic (30M035) et de RT-PCR sur LightCycler (30M543). (20210107 BV RUN2 21GM000806)

Résultats.

FAM	21GM000467	21GM0005178	21GM0005202
Cq 1	22,61	15,83	Négatif
Cq2	22,37	17,06	Négatif
Cq3	22,07	16,39	Négatif
Moyenne	22,35	16,42666667	/
Ecart Type	0,270554985	0,615819238	/
CV	0,012105368	0,037488996	/

HEX	21GM000467	21GM0005178	21GM0005202
Cq 1	21,11	14,73	Négatif
Cq2	21,16	15,90	Négatif
Cq3	22,53	14,80	Négatif
Moyenne	21,6	15,14333333	/
Ecart Type	0,805791536	0,656226587	/
CV	0,037305164	0,043334355	/

Cy5	21GM000467	21GM0005178	21GM0005202
Cq 1	22,72	21,21	28,65
Cq2	22,60	21,57	29,00
Cq3	22,45	20,64	28,35
Moyenne	22,59	21,14	28,66666667
Ecart Type	0,135277493	0,468934964	0,325320355
CV	0,005988379	0,022182354	0,011348384

4. Practical application to molecular diagnostics

qPCR verification: reproductibility

Méthodologie et échantillons. 6 échantillons positifs pour SARS-CoV-2 et 3 échantillons négatifs pour SARS-CoV ont été analysés en simplicat au cours d'un run extraction Chemagic et 3 runs de RT-PCR : jour, TL, LightCycler480 différents.

Résultats.

FAM	21GM000408	21GM000281	21GM000443	21GM000282	21GM000394	21GM000407	21GM000508	21GM000509	21GM000510
Jour 1	16,55	15,46	19,67	28,37	27,67	28,77	Négatif	Négatif	Négatif
Jour 2	16,56	15,06	19,84	28,57	27,90	28,73	Négatif	Négatif	Négatif
Jour 3	16,08	14,96	19,60	28,41	27,93	28,16	Négatif	Négatif	Négatif
Moyenne	16,39666667	15,16	19,70333333	28,45	27,83333333	28,55333333	/	/	/
Ecart Type	0,274286954	0,264575131	0,123423391	0,105830052	0,142243922	0,34122329	/	/	/
CV	0,016728214	0,017452185	0,006264087	0,003719861	0,00511056	0,011950384	/	/	

HEX	21GM000408	21GM000281	21GM000443	21GM000282	21GM000394	21GM000407	21GM000508	21GM000509	21GM000510
Jour 1	16,25	14,97	20,27	27,97	26,36	29,24	Négatif	Négatif	Négatif
Jour 2	16,25	14,01	19,87	27,99	26,32	28,79	Négatif	Négatif	Négatif
Jour 3	16,12	14	19,95	27,97	26,48	28,27	Négatif	Négatif	Négatif
Moyenne	16,20666667	14,32666667	20,03	27,97666667	26,38666667	28,76666667	/	/	/
Ecart Type	0,075055535	0,557165445	0,211660105	0,011547005	0,08326664	0,48542078	/	/	/
CV	0,004631152	0,038890096	0,010567155	0,000412737	0,003155633	0,016874419	/	/	

Cy5	21GM000408	21GM000281	21GM000443	21GM000282	21GM000394	21GM000407	21GM000508	21GM000509	21GM000510
Jour 1	22,83	25,28	23,78	22,67	24,58	25,01	21,68	25,66	26,39
Jour 2	22,99	24,89	23,8	22,96	25,54	24,96	21,54	28,19	27,89
Jour 3	22,74	24,57	23,51	22,82	25,2	24,61	19,06	25,62	27,68
Moyenne	22,85333333	24,91333333	23,69666667	22,81666667	25,10666667	24,86	20,76	26,49	27,32
Ecart Type	0,126622799	0,355574652	0,161967075	0,145028733	0,486757982	0,217944947	1,473906374	1,472379027	0,812219182
CV	0,005540671	0,014272464	0,006835015	0,006356263	0,019387599	0,008766892	0,070997417	0,055582447	0,029729838

Conclusion. 100% de concordance dans les résultats inter-run.

4. Practical application to molecular diagnostics

qPCR verification: accuracy

3. Justesse / exactitude (+ calcul d'incertitude de mesure si justifié)

Est mesurée par le biais entre la moyenne d'un nombre infini de valeurs mesurées répétées et la valeur de référence (μ).

$$\text{Biais} = \text{moyenne des résultats (M)} - \text{valeur de référence } (\mu)$$

Exactitude (erreur totale) = Justesse (erreur systématique) + fidélité (Erreur aléatoire)

=


Biais + Ecart-type

Objectif. Confirmation de la présence de SARS-CoV-2 via un QC externe QCMD.

Méthodologie et échantillons. 15 QC SARS-CoV-2 ont été extraits et analysés. Les résultats ont été envoyés à l'organisme organisateur.

4. Practical application to molecular diagnostics

qPCR verification: accuracy

Individual Report		QCMD 2020 SARS-CoV-2 EQA Programme				
Catalogue Code: QAV204215	Ref Code: SCV2_20	Challenge: C2	Analysis Type: Qualitative	Dataset: 401539	Report UID: 1197/401539/3130	Laboratory BE034

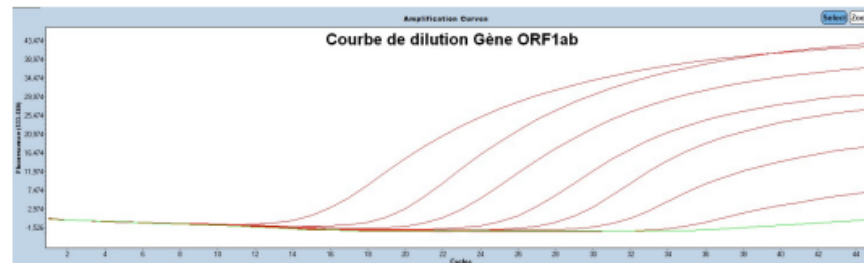
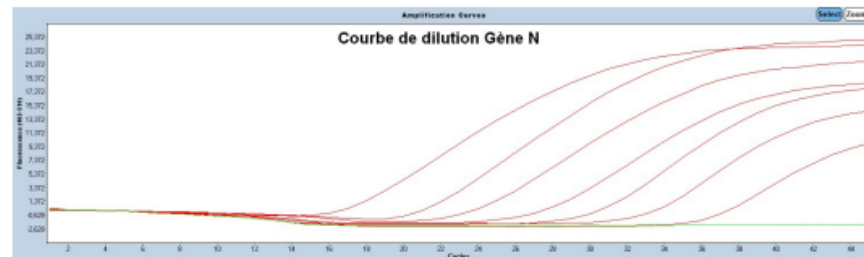
Core Panel Members Results

Sample Code	Qualitative Results			Your Quantitative Data (for information only) ^[3]		
	Percentage Correct (All) ^[4]	Your Result ^[5]	Detection Score ^[6]	Reported Value	Unitage	Cycle Threshold
SCV2_101C2-01	97.8	Negative	0		N/A	-
SCV2_101C2-02	98.7	Positive	0		N/A	32.79
SCV2_101C2-03	93.1	Positive	0		N/A	35.98
SCV2_101C2-04	93.6	Positive	0		N/A	35.65
SCV2_101C2-05	99.1	Negative	0		N/A	-

4. Practical application to molecular diagnostics

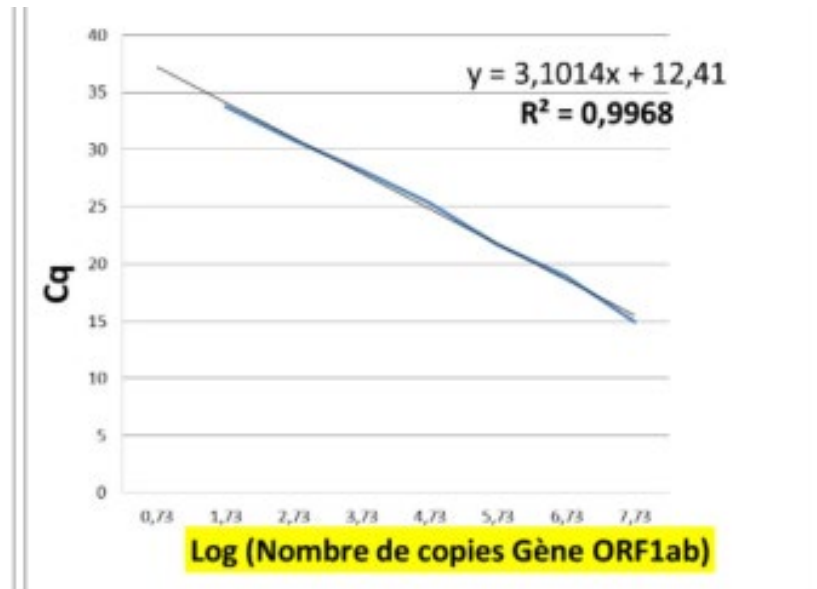
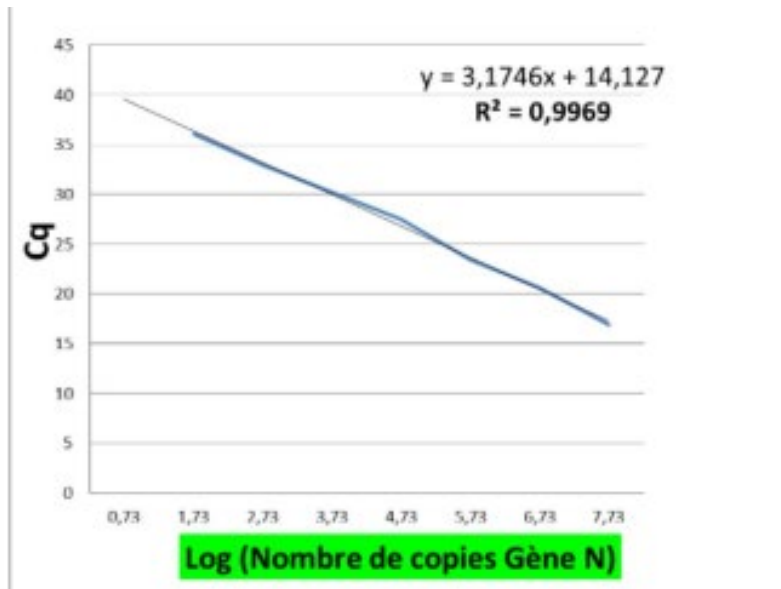
qPCR verification: linearity

N ^{bre} Copies	Log	Cq FAM	Cq Hex
5,37E+07	7,73	16,91	14,96
5,37E+06	6,73	20,63	18,91
5,37E+05	5,73	23,46	21,7
5,37E+04	4,73	27,51	25,33
5,37E+03	3,73	30,26	28,17
5,37E+02	2,73	32,95	30,85
5,37E+01	1,73	36,06	33,79
5,37E+00	0,73		



4. Practical application to molecular diagnostics

qPCR verification: linearity



Statistical methods in validation reports and QC

Important messages

- **Impact of IVD-R and new ISO 15189:2022 version**
- **Validation vs Verification**
- **Statistics are important for validation/verification**
- **Statistical QC can be applied to molecular diagnostics**

Thanks for your participation !

jeremie.gras@ipg.be



Thanks for your attention !

jeremie.gras@ipg.be



Statistical methods in validation reports and QC

Jeremie Gras, M.D.
Medical Director, IPG
MB&C Course 2024



Institut de Pathologie et de Génétique
25, avenue Georges Lemaître B6041
Gosselies
Tél : +32 (0)71 47 30 47
accueil@ipg.be

