Quality Control of Qualitative Tests for Medical Laboratories

Introduction to the book

Paulo Pereira, Ph.D.

ISBN 978-989-20-9857-9

Quality Control of Qualitative Tests
for Medical Laboratories

What is required, and what is not, in the ISO standards? Which are the most significant sources of uncertainty? What is the similarity and difference between "Uncertainty Approach," and "Error Approach"? Which models do we use to compute both methodologies? And which models to determine conditional accuracy, delta values, and seronegative window period? Which are the best models to compute the agreement of binary results? How do we identify "the best" cutoff point? How do we control the performance of the qualitative results in daily routine? More than 20 examples based on real-world data are presented. The book includes several cases of immunoassays and NAT for screening in virology, ABO blood test, HLA typing, and karyotype tests. The statistical quality control tools applied to the examples are generic; they can be used in most of the qualitative tests.

Paulo Pereira was born on March 10, 1972, in Lisbon, Portugal. He received his Ph.D. from the Catholic University of Portugal (Biotechnology, specialization in Microbiology). Dr. Pereira is a Senior Researcher and the Head of the Research & Development Department of the Portuguese Institute of Blood and Transplantation (IPST). He has been recruited as a Quality and Laboratory Expert for seminars and professional laboratory meetings throughout Europe, South America, and Africa. He has 25+ years of experience in a medical laboratory, having held key scientific leadership roles: 9+ years as a Medical Technician, 15+ years as a Researcher, and 5+ years as a Consultant of a Metrology Laboratory. He has been work for 20+ years as a Consultant and Auditor of Quality Management Systems and Technical Requirements (ISO 9001, ISO/IEC 17025, ISO 15189, and ISO 13485). He has 15+ years of experience as a Quality Manager and the National Coordinator of Quality Assurance in the IPST. He has 15+ years of experience as a University Professor. He has authored several peer-review scientific and technical articles, and several indexed book chapters. He is also a reviewer of several scientific and technical articles and a member of Editorial Boards. Dr. Pereira is a Technical Expert on CLSI Document Development Committee on EP12.

ISBN 978-989-20-9857-9

Why publish a book called "Quality control of qualitative tests for medical laboratories"?

- Address the need for a book dedicated to quality control of qualitative tests
- The is a book written primarily for the laboratorian and aims to substantiate the selection of the best statistical tools considering the intended use of the qualitative tests' results (fitness for purpose)
- The purpose of the book is to answer most of qualitative tests QC questions in a three-pronged vision: the statistical, the clinical and the regulatory vision

- The book seeks to answer questions important to laboratory practice such as:

  – What is required, and what is not, in the ISO standards?

  – Which are the most significant sources of uncertainty?

  – What is the similarity and difference between “Uncertainty Approach,” and “Error Approach”?

  – Which models do we use to compute both methodologies?

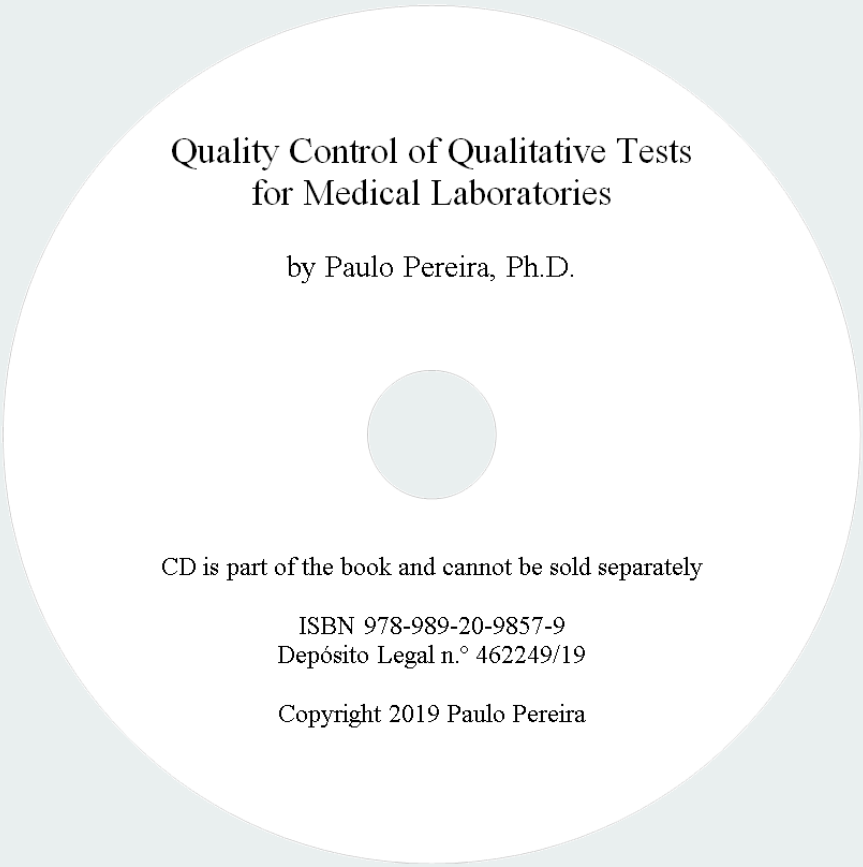
  – And which models to determine conditional accuracy, delta values, and seronegative window period?

  – Which are the best models to compute the agreement of binary results?

  – How do we identify “the best” cutoff point?

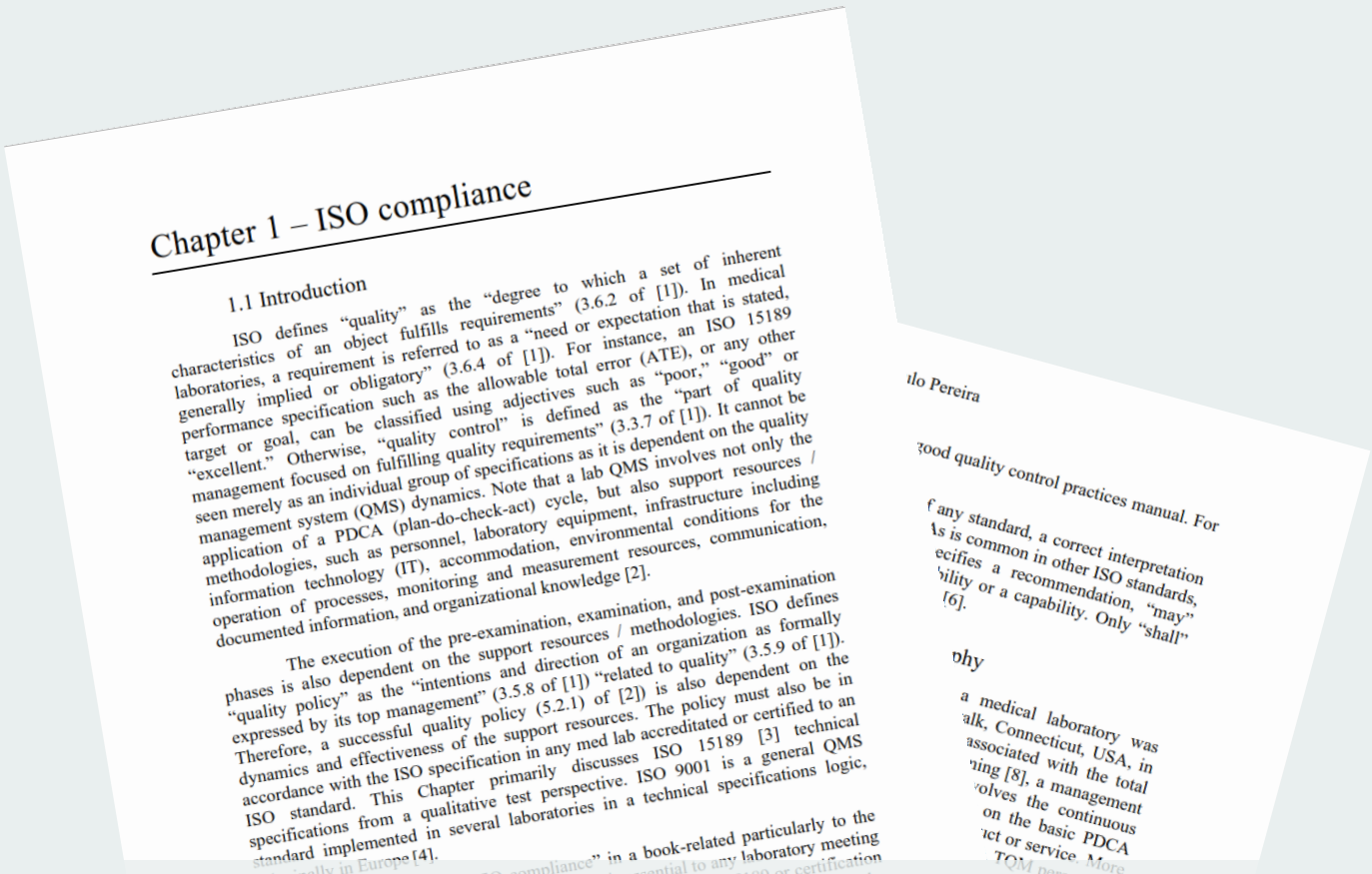
  – How do we control the performance of the qualitative results in daily routine?

- More than 20 examples based on real-world data are presented
- The book includes several cases of immunoassays and NAT for screening in virology, ABO blood test, HLA typing, and karyotype tests
- The statistical quality control tools applied to the examples are generic; they can be used in most of the qualitative tests
- Approx. 200 pages printed on coated paper (couché) 90 grams; cover printed on 170 gram coated paper with soft-touch plastic coating; 2mm hard card cover

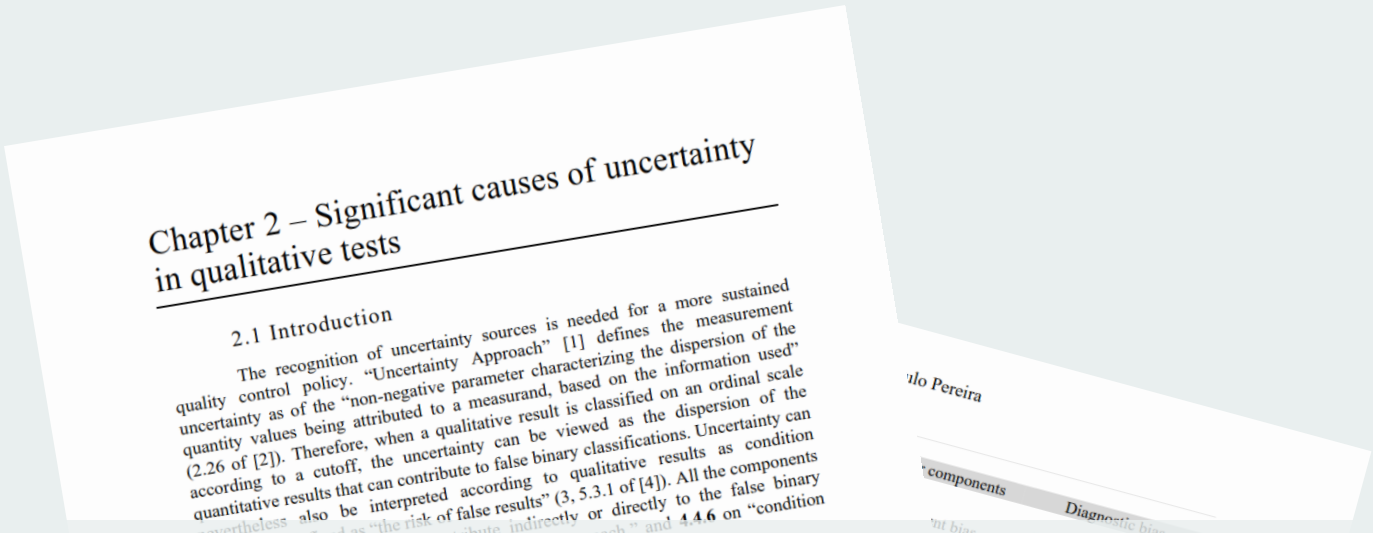


Why include a CD with spreadsheets?

- For a more natural comprehension of the approaches
- Facilitate the understanding of theory based on practice
- A practical way to demonstrate the case studies included in the book
- The laboratorian can easily replicate the models for his practice
- All the computations can be done using a conventional computer spreadsheet
- Excel® (Microsoft®, Redmond, Washington, USA) is immediately recognized as very intuitive software for laboratorian
- Readers will receive free updates to the spreadsheet package



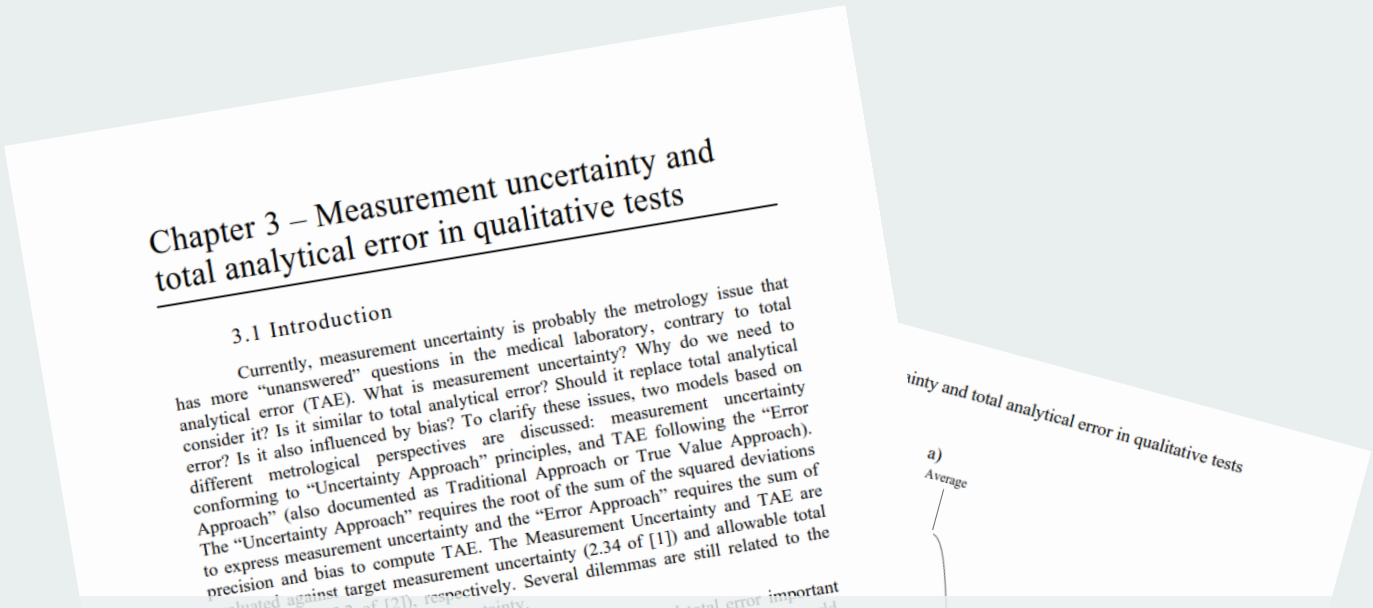
- **Chapter 1 – ISO compliance** introduces mainly ISO 15189 for the accreditation of medical laboratory methods or tests
- For a consistent application of this global standard, the laboratorian must understand its specifications
- We have discussed the use of most of its technical requirements that involve the selection, verification, validation, measurement uncertainty, internal quality control, and external quality assessment / proficiency testing (EQA / PT) of qualitative results
- Moreover, we have crossed ISO 15189 with ISO 9001 requirements for a more natural interpretation of this guideline, which is oriented to a generic implementation of a quality management system
- How do we meet the referred ISO claims? See the following chapters for suggested methodologies


Chapter 2 – Significant causes of uncertainty in qualitative tests

2.1 Introduction

The recognition of uncertainty sources is needed for a more sustained quality control policy. "Uncertainty Approach" [1] defines the measurement uncertainty as of the "non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used" (2.26 of [2]). Therefore, when a qualitative result is classified on an ordinal scale according to a cutoff, the uncertainty can be viewed as the dispersion of the quantitative results that can contribute to false binary classifications. Uncertainty can

- **Chapter 2 – Significant causes of uncertainty in qualitative tests** discusses the main sources of error that can cause untrue binary results
- As the test methodology is essential to recognize the most common analytical causes of failure, we have presented a brief overview of qualitative test design
- The impact of the analytical error on the cutoff trueness is discussed, as well as the effect of the analytical error on the accuracy of the classification of binary results
- The importance of the "gray zone" and the associated trinary classification to minimize the impact of analytical error in the results is debated
- The biased results due to biological factors are presented with a focus on the seroconversion window period
- The contribution of other possible sources of bias to the lack of representativeness of patients' samples is also pondered.
- The impact of interferences in bias is discussed
- This debate is important for a better focus on the use of the quality control tools that allow us to see what is and what is not measurable (limitation of the studies)



- **Chapter 3 – Measurement uncertainty and total analytical error in qualitative methods** introduces both the Uncertainty Approach" and the "Error Approach"
- The challenge is to introduce the laboratorian to the similarities and differences of the visions, wherein empirical models are considered for both visions
- While not ignoring the usefulness of the modular models to the manufacturer, they are not discussed further here since they are not meant to be used in medical laboratory practice
- The models presented are based on recognized protocols in med lab requiring data from single-laboratory validation, interlaboratory comparisons or EQA /PT
- The importance of the metrological traceability of the results is considered
- Compliance assessment is associated with the empirical estimate of the "gray zone" and the limit of detection (LoD)
- The evaluation of analyte concentrations near the cutoff is presented as a complementary tool to estimate an identical zone

Chapter 4 – Performance of binary classification tests

4.1 Introduction

Performance of binary classification (true / false, positive / negative) test can be determined by several statistical measures such as sensitivity and specificity. Diagnostic accuracy is one example of a statistical model to compute binary results performance. In this case, the "condition" is the disease. It is defined as "the ability of a diagnostic test to discriminate between diseased and non-diseased subjects or between two or more clinical states" [1]. It is probably the most recognized statistical measurement in medical laboratories to determine the performance of a qualitative test for the detection of infections. Nevertheless, some points require clarification for a more robust application. What is the added value of the confidence interval? What can be inferred to the population? What are the limitations of the evaluation?

A test result is considered "positive" if it is characteristic of a positive condition and "negative" if characteristic of negative condition populations. Although usually test results are negative when lower than the cutoff, and positive when equal or higher, this is not always the case, as for example when using a competitive method. For example, to diagnose hypothyroidism, T_4 is "positive" (characteristic of patients) if inferior to cutoff.

The statistical principles are based on Bayesian probability models. Typically, the methodology uses a 2x2 contingency table to compute the condition sensitivity and specificity using positive condition and negative condition samples, respectively. For an easier understanding, the Chapter mainly refers to the case of individuals with the disease (positive condition) and individuals without the disease (negative condition) with the performance interpreted as "diagnostic accuracy." The same logic can be applied to any other true condition (also referred to as target condition or condition of interest) as determined by the accuracy criteria. Examples include any other identifiable condition within an individual, such as a antigen (HLA) type, health status, blood groups, a karyotype, human leukocyte

lo Pereira

...etrology (VIM) defines "measurement ... characterizing the dispersion of the ...nd, based on the information used" ...r qualitative values, metrologically ...n a risk-based thinking perspective ...nsider "the effect of uncertainty" ...ision stage), such as method ...at complex challenge, as the ...tainty just include modeling ... do not integrate alternative ...he results of the qualitative ... "diagnostic uncertainty" ...s book is quality control ...n suggested by Pereira ...sults." This definition ...ne the "diagnostic ... a more natural

- **Chapter 4 – Performance of binary classification tests** is based on condition accuracy, probably the most well-known methodology for validating qualitative results
- In this chapter, we introduce the basis of the statistics concepts applied and discuss the importance of the samples to the robustness of the estimates
- We have used 2x2 contingency tables, followed by a discussion about the value of the analysis of the numerical data to distinguish between two or more tests with identical condition sensitivity and specificity
- The concept of "condition uncertainty" is introduced, analogous to the "measurement uncertainty" of quantitative dimensions
- The window period is presented using a binary and trinary

Chapter 5 – Agreement of binary classification tests

5.1 Introduction

Every so often, samples with the known condition for evaluating qualitative tests are not available, especially those intended to detect rare conditions, such as diseases. As an alternative, samples with known results in a comparator test are suggested (10.2 of [1]). A disadvantage is that it is not possible to compute condition accuracy rates (see **Chapter 4**). Sometimes the agreement is misunderstood with condition accuracy such as diagnostic accuracy. In this case, a bad practice should be immediately recognized, as condition accuracy rates require evidenced-based condition information as diagnostic info. Condition sensitivity and specificity cannot be determined just by the results of a test that is not a "gold standard" (see **4.3.1**). Misinterpretation of the accuracy of results of a new test could, therefore, happen, and the correct use of the binary may be compromised. If the positive condition and negative condition samples are not available, one could ask a few questions. How do we determine the agreement of binary results? Does the confidence interval add value to the evaluation? Can the agreement be inferred to the population? What are the limitations of the study?

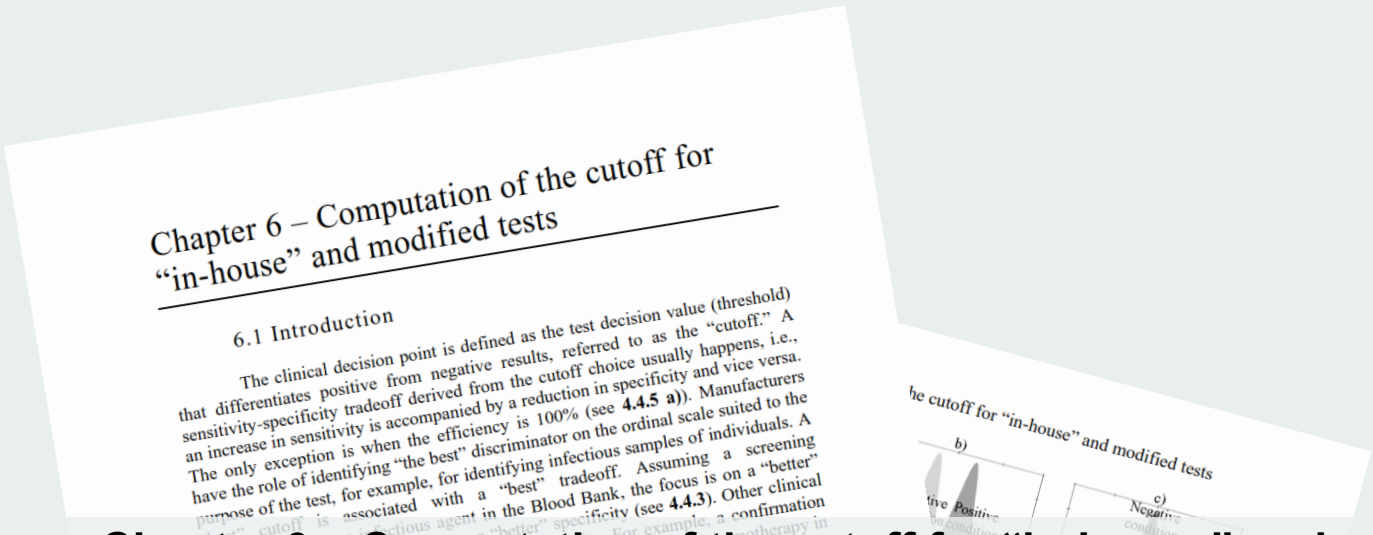
The Chapter presents a set of good practices to guarantee reliable agreement evaluations. These practices are suited to support the robust use of agreement methods. Real-world data is used to illustrate the models.

5.2 Principles

The purpose of the agreement of binary results is to evaluate a test based on the agreement of the results of a candidate and a comparative test. The comparator results are assumed as those with the highest probability of being true. Thus, the accuracy of the study depends on the comparative test performance. This test should not be confused with a "gold standard." Thus, the comparator should show at least state-of-the-art performance, i.e., a demonstrated top performance of condition accuracy. This information can be collected in a review of the literature for most of the assays. The candidate test performance decreases proportionally with the number of discrepant results against the comparative assay. Then, the selection of the comparator is critical to the reliability and consistency of the evaluation. For example, the U.S. Food and Drug Administration (FDA) ranking (V.C.1 of [2]) classifies the quality of a comparative method as follows:

- type A as "a quantitative reference method (…) with the appropriate cutoff value for the positive and negative results";
- type B as "a qualitative reference method (…)";

- **Chapter 5 – Agreement of binary classification tests** is intended to lead the reader to validation where samples with a true condition are unavailable
- Since the consistency of the results is dependent on the comparative test performance, its selection should be applied uniquely if the condition is unknown


Chapter 6 – Computation of the cutoff for "in-house" and modified tests

6.1 Introduction

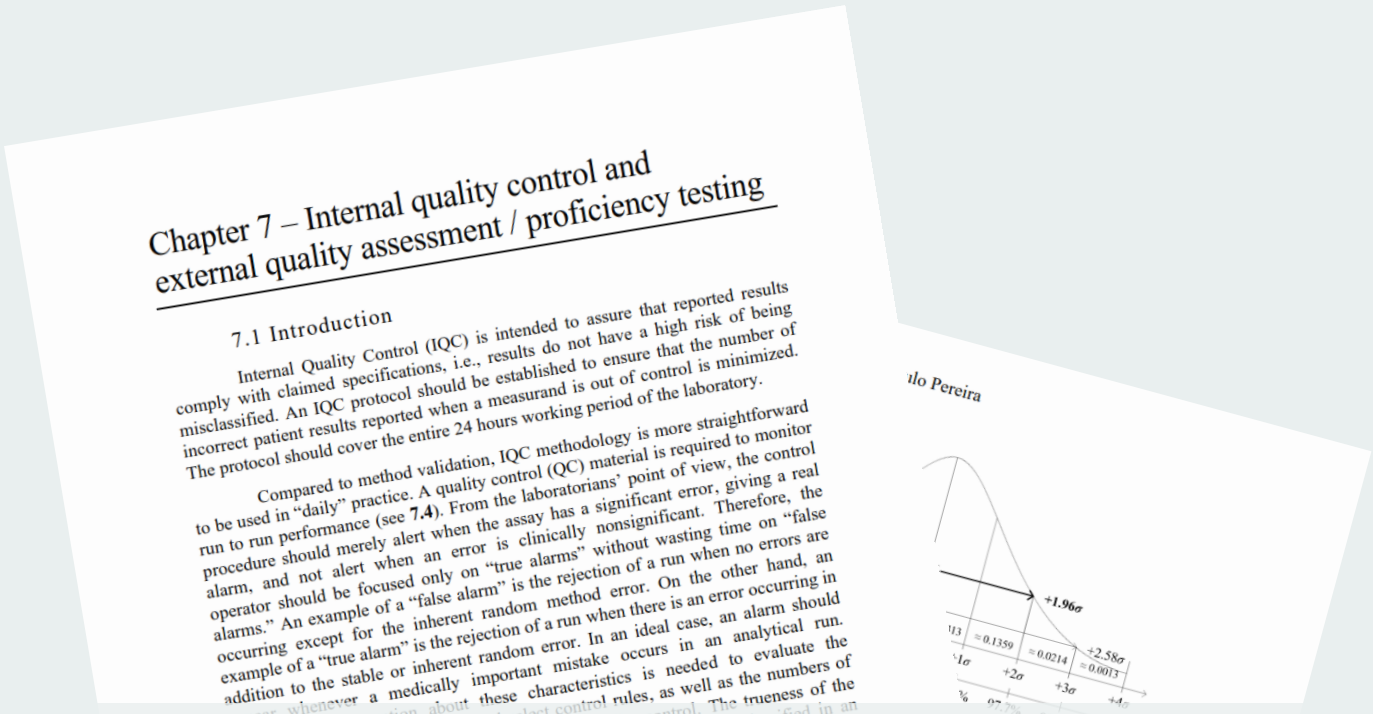
The clinical decision point is defined as the test decision value (threshold) that differentiates positive from negative results, referred to as the "cutoff." A sensitivity-specificity tradeoff derived from the cutoff choice usually happens, i.e., an increase in sensitivity is accompanied by a reduction in specificity and vice versa. The only exception is when the efficiency is 100% (see **4.4.5 a)**). Manufacturers have the role of identifying "the best" discriminator on the ordinal scale suited to the purpose of the test, for example, for identifying infectious samples of individuals. A

he cutoff for "in-house" and modified tests

b)

c)

- **Chapter 6 – Computation of the cutoff for "in-house" and modified tests**, as the title refers, applies solely to tests prepared in the laboratory requiring cutoff determination
- The "realism" of the cutoff does not depend only on the samples but also on the intended use of the results
- Usually, false-positive results are better accepted than false-negative ones
- The computation of the cutoff by the receiver operating characteristic curve (ROC) is discussed
- Although we have tried to use the most accessible language, it is probably the most complex statistical model presented in this book
- However, its principle is simple: it provides the various condition sensitivities and specificities for all the possible cutoff points
- The laboratorian selects the point that meets the requirements related to the intended use of the results, i.e., according to the clinical application
- An area ranking allows the classification of the detection capability of the test for a certain cutoff


Chapter 7 – Internal quality control and external quality assessment / proficiency testing

7.1 Introduction

Internal Quality Control (IQC) is intended to assure that reported results comply with claimed specifications, i.e., results do not have a high risk of being misclassified. An IQC protocol should be established to ensure that the number of incorrect patient results reported when a measurand is out of control is minimized. The protocol should cover the entire 24 hours working period of the laboratory.

Compared to method validation, IQC methodology is more straightforward to be used in "daily" practice. A quality control (QC) material is required to monitor run to run performance (see **7.4**). From the laboratorians' point of view, the control procedure should merely alert when the assay has a significant error, giving a real alarm, and not alert when an error is clinically nonsignificant. Therefore, the operator should be focused only on "true alarms" without wasting time on "false alarms." An example of a "false alarm" is the rejection of a run when no errors are occurring except for the inherent random method error. On the other hand, an example of a "true alarm" is the rejection of a run when there is an error occurring in addition to the stable or inherent random error. In an ideal case, an alarm should appear whenever a medically important mistake occurs in an analytical run. Information about these characteristics is needed to evaluate the control rules, as well as the numbers of

lo Pereira

- **Chapter 7 – Internal quality control and external quality assessment / proficiency testing** debate models suitable for qualitative tests
- The internal quality control principles are discussed to aid the selection of the best designs based on a qualitative logic
- Demystification of control rules in qualitative tests statistically and clinically supported
- Novel approaches to compute sigma metrics in qualitative tests
- The DPMO-derived and SE_{crit}-derived sigma metrics express the capability of tests to meet the specifications
- Models are presented for variables using numerical results (ordinal tests), and an application to monitor "pure" qualitative results (nominal tests)
- Both methodologies are intended to <u>control the loss of sensitivity</u> in the qualitative tests
- EQA /PT is introduced

Quality Control of Qualitative Tests for Medical Laboratories

Introduction to the spreadsheets

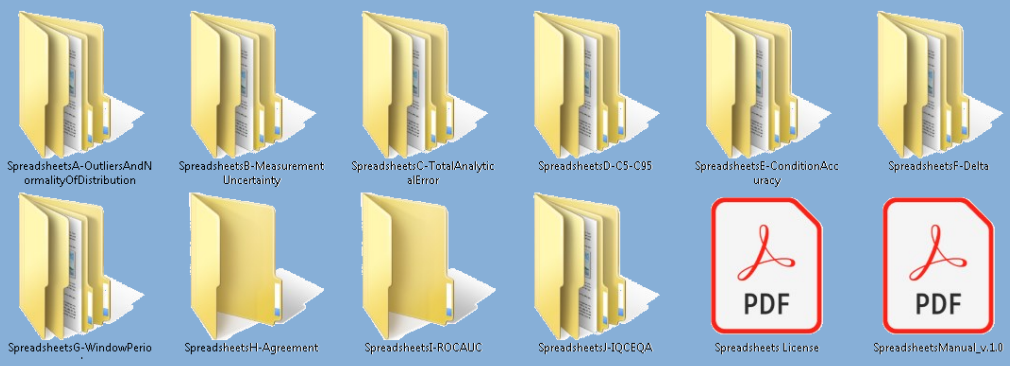
Paulo Pereira, Ph.D.

ISBN 978-989-20-9857-9

Excel® (Microsoft®, Redmond, Washington, USA) spreadsheets presents several advantages:

- Study Cases' Demonstration
- Office Integration
- Formula and Calculation Features
- Use Conditional Formatting
- Bring Data Together
- Easy and Effective Comparisons
- Microsoft Excel Mobile & iPad Apps
- Expansion to Book Contents
- Data Organization
- Build Great Charts
- Help Identify Trends
- Online Access
- Powerful Analysis of Large Amounts of Data
- Third-Party Support

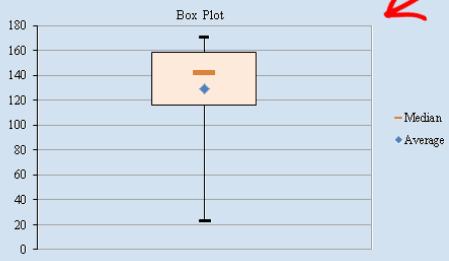
- The CD includes more than 25 spreadsheets to demonstrate all the examples included in the book
- Supplemental study cases are included
- Also Includes empty spreadsheets for med lab data computation
- Spreadsheets are designed and verified using harmonized models



A. Outliers and Normality Evaluation

Verification of outliers

	n >= 3	Grubbs' Test				Box-plot	
no.	Data	Min =	23.000	Max =	171.000	1st Quartile =	116
1	120.000	Next to min =	104	Next to max =	162	Min =	23
2	142.000					Average =	129.6364
3	112.000	Average =	129.6364	alpha =	0.05	Median =	142.000
4	162.000	*s* =	41.68758	*n* =	11	Max =	171
5	155.000	G =	2.557989	Sig value =	0.004545	3rd Quartile =	158.5
6	148.000			*df* =	9		
7	23.000			*t*-crit =	3.309517		
8	171.000			G-Crit =	2.233908		
9	104.000						
10	127.000	Suspected outliers:	**Yes**				
11	162.000						



	n >= 20	Skewness test		Kurtosis test		D'Agostino-Pearson Omnibus Test	
no.	Data	*n* =	24	*n* =	24	Stat =	0.904255863
1	34.000	alpha =	0.05	alpha =	0.05	*p*-value =	**0.64**
2	56.000	Skewness =	0.195701	Kurtosis =	-0.78551	Result:	**Accept Normality**
3	39.000	s.e. =	0.472261	s.e. =	0.917777		
4	71.000	Test stat =	0.414391	Test stat =	-0.85588		
5	84.000	*p*-value =	0.339294	*p*-value =	0.196031		
6	92.000	Lower =	-0.72991	Lower =	-2.58432		
7	44.000	Upper =	1.121315	Upper =	1.0133		

Data entry made easy using cells marked in yellow

Verification of normal distribution

B. Measurement Uncertainty (Empirical Approaches)

Single-laboratory validation approach using short-term data

Test:	anti-HCV		Units:	S/CO

Technician:	PPereira	PPereira	PPereira	PPereira	PPereira
Date:	March 24, 2019	March 25, 2019	March 26, 2019	March 27, 2019	March 28, 2019
	Run 1	Run 2	Run 3	Run 4	Run 5
Rep 1	1.020	1.080	1.000	1.010	1.030
Rep 2	1.100	0.950	1.030	0.970	1.020
Rep 3	0.910	0.940	0.970	1.030	0.900
Rep 4	0.850	0.990	0.950	1.070	1.000
Rep 5	0.940	1.000	0.940	0.950	1.040
n =	5	5	5	5	5
df =	4	4	4	4	4
df =	20				
Sum =	4.82	4.96	4.89	5.03	4.99
Mean =	0.964	0.992	0.978	1.006	0.998
Variance =	0.00953	0.00307	0.00137	0.00228	0.00322
SSW =	0.07788				

Source of variation	SS	DF	MS	n_0	V_B
Between-run =	0.005576	4	0.001394	5	0
Within-run =	0.07788	20	0.003894		
Total =	0.083456				

s_r =	0.062401923
s_I =	0
s_{Rw} =	**0.062401923**

Test:	anti-HCV	Units:	S/CO

Mean	Cert.value	*bias*	s_{bias}	n	U_{c_ref}	k	u_{c_ref}	u_{bias}
0.992	0.980	1.22%	2.1%	5	0	1.96	0.0%	**1.5%**

Run no.	Technician	Date	Replicate 1	Replicate 2	Replicate 3	Replicate 4	Replicate 5	Mean	*bias*	Relative *bias*
1	PP	March 30, 2018	0.890	0.920	0.970	0.980	1.030	0.958	-0.022	-2.2%
2	PP	March 31, 2018	1.010	1.080	0.920	0.930	1.090	1.006	0.026	2.7%
3	PP	April 1, 2018	1.060	0.940	1.050	1.020	0.980	1.010	0.030	3.1%
4	PP	April 2, 2018	1.100	0.920	0.950	0.960	1.050	0.996	0.016	1.6%
5	PP	April 3, 2018	1.060	1.060	1.040	0.880	0.910	0.990	0.010	1.0%

Test:	anti-HCV	Units:	S/CO

s_{Rw}	u_{bias}	u	k	U
6.3%	1.5%	6.5%	2.00	**13.0%**

Single-laboratory validation approach using long-term data

Test:	anti-HCV	Units:	S/CO

s_{Rw}	u_{bias}	u	k	U
9.0%	1.5%	9.2%	2.00	**18.3%**

			Technician:	PPereira		Technician:	PPereira					
			Date:	March 24, 2019		Date:	March 24, 2019					
			Run 1			Run 2				Run 1	Run 2	
Run no.	Techn.	Date	Result 1	Result 2	Mean Run 1	Result 1	Result 2	Mean Run 2	Daily Mean	(Rep 1 - Rep 2)^2	(Rep 1 - Rep 2)^2	(Mean Run 1 - Mean Run 2)^2
1	PP	March 4, 2018	1.090	0.990	1.040	1.070	1.010	1.040	1.040	0.01	0.0036	0
2	PP	March 5, 2018	0.980	0.890	0.935	1.010	1.040	1.025	0.980	0.0081	0.0009	0.0081
3	PP	March 6, 2018	1.060	0.890	0.975	1.040	0.880	0.960	0.968	0.0289	0.0256	0.000225
4	PP	March 7, 2018	0.860	1.070	0.965	1.010	1.080	1.045	1.005	0.0441	0.0049	0.0064
5	PP	March 8, 2018	0.880	1.000	0.940	0.950	1.040	0.995	0.968	0.0144	0.0081	0.003025
6	PP	March 9, 2018	0.960	0.890	0.925	0.890	0.880	0.885	0.905	0.0049	0.0001	0.0016
7	PP	March 10, 2018	1.100	0.980	1.040	1.060	0.960	1.010	1.025	0.0144	0.01	0.0009
8	PP	March 11, 2018	0.960	0.890	0.925	0.880	1.030	0.955	0.940	0.0049	0.0225	0.0009
9	PP	March 12, 2018	0.900	0.960	0.930	0.850	1.100	0.975	0.953	0.0036	0.0625	0.002025
10	PP	March 13, 2018	0.940	0.850	0.895	0.900	1.040	0.970	0.933	0.0081	0.0196	0.005625
11	PP	March 14, 2018	1.090	0.960	1.025	1.000	1.060	1.030	1.028	0.0169	0.0036	2.5E-05
12	PP	March 15, 2018	0.930	0.890	0.910	0.950	1.090	1.020	0.965	0.0016	0.0196	0.0121
13	PP	March 16, 2018	1.020	0.890	0.955	0.860	0.990	0.925	0.940	0.0169	0.0169	0.0009
14	PP	March 17, 2018	1.040	0.910	0.975	0.930	1.050	0.990	0.983	0.0169	0.0144	0.000225
15	PP	March 18, 2018	1.040	0.900	0.970	1.080	1.010	1.045	1.008	0.0196	0.0049	0.005625
16	PP	March 19, 2018	1.090	0.920	1.005	1.040	0.890	0.965	0.985	0.0289	0.0225	0.0016
17	PP	March 20, 2018	1.000	0.990	0.995	1.080	1.090	1.085	1.040	0.0001	0.0001	0.0081
18	PP	March 21, 2018	0.940	1.080	1.010	1.080	1.040	1.060	1.035	0.0196	0.0016	0.0025
19	PP	March 22, 2018	0.940	1.040	0.990	1.070	1.100	1.085	1.038	0.01	0.0009	0.009025
20	PP	March 23, 2018	0.870	0.900	0.885	1.040	0.850	0.945	0.915	0.0009	0.0361	0.0036
									0.983	0.2728	0.2784	0.0725
									Mean	Σ	Σ	Σ

n = 20		ME =	0.00689	Repeatability (s_r)=	0.083006024	8.4%
R = 19.65		MR =	0.003625	Between day SD (s_{dd})=	0.031188097	3.2%
A = 0.042573466		MD =	0.007515789	Between run SD (s_{ir})=	0	0.0%
B = 0.043346827		T =	74	**s_{Rw} =**	**0.088671852**	**9.0%**

Detailed presentation of statistical calculations for a more natural understanding of models

B. Measurement Uncertainty (Empirical Approaches)

Interlaboratory comparisons

Using intercomparison data

Test:	anti-HCV	Units:	S/CO			
Reference result	Lab result	Deviation	s_R		No. of ref labs	Excluded
0.980	0.992	1.2%	0.180	18.4%	2	0

s_{Rw}	u_{bias}	u	k	U
18.4%	0.0%	18.4%	2.00	**36.7%**

External quality assessment / proficiency testing

Test:	anti-HCV	Units:	S/CO		
Bias average	s_R	No. of labs average	RMS_{bias}	u_{c_ref}	u_{bias}
-2.4%	12.2%	9	6.8%	4.0%	7.9%

Exercise	Ref result	Lab result	*bias*	Bias^2	s_{R_group}	No. of labs
1	2.2	2.1	-4.5%	0.207%	12.1%	7
2	2.4	2.2	-8.3%	0.694%	13.1%	11
3	2.6	2.4	-7.7%	0.592%	11.4%	9
4	1.9	2.1	10.5%	1.108%	12.6%	12
5	2.2	2.2	0.0%	0.000%	11.8%	8
6	2.3	2.2	-4.3%	0.189%	12.4%	9

Using EQA / PT data for bias uncertainty estimate

Test:	anti-HCV	Units:	S/CO	
s_{Rw}	u_{bias}	u	k	U
9.0%	7.9%	12.0%	2.00	**24.0%**

Combination with within-laboratory reproducibility standard deviation

Computation of coverage factors derived from effective degrees of freedom

Coverage factors (k) derived from effective degrees of freedom (ν_{eff})								
Factor	c_i	x_i	Unit	Type	Distribution	Divisor	$u(x_i)$	df^*
s_{RW}	1	0.976	%	A	Gaussian	2	6.3	20
u_b	1	0.992	%	A	Gaussian	2	1.5	20

	$u_c(y)$	ν_{eff}
Combined uncertainty (RSS method) =	6.47611	22.26031
Expansion coefficient (k) =	2.131106	
Expanded uncertainty (U) =	13.80128	%

Most accurate k value

* $df = n - 1$

* Welch-Satterthwaite formula

B. Measurement Uncertainty (Empirical Approaches)

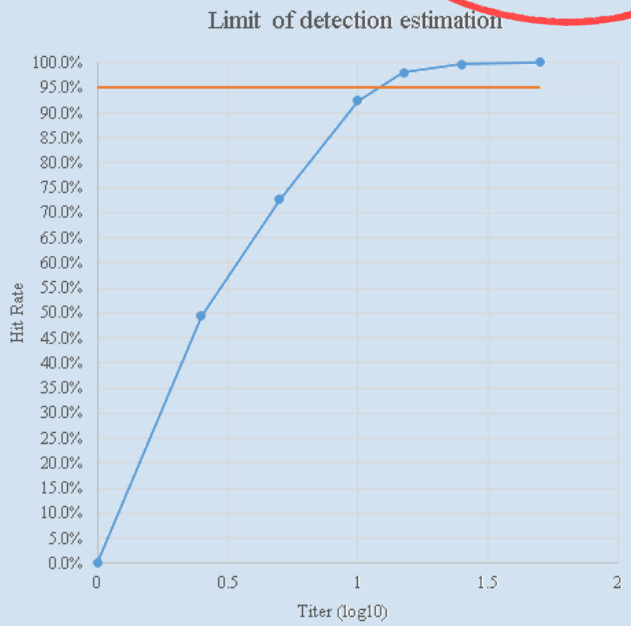
Limit of detection (LoD) in molecular biology tests (probit regression model)

Test:	RT-PCR RNA SARS-Cov-2
Units:	copies / reaction
Technician:	PPereira
Date:	February 4, 2020

The percentage proportion of positive results

Input titer	log_10	No. of valid replicates	No. of positives	Hit rate	Logit(P)	**Probit(P)**
50	1.699	252	252	100.0%		
25	1.398	252	251	99.6%	5.5254529	7.654759
15	1.176	251	246	98.0%	3.8958936	7.055397
10	1.000	252	233	92.5%	2.5065995	6.436734
5	0.699	252	183	72.6%	0.9753796	5.601332
2.5	0.398	251	124	49.4%	-0.023906	4.98502
0	0.000	250	0	0.0%		

Slope:	2.72147
Intercept:	3.804283
Test value:	6.64
Log_10:	1.04198
LC95:	**11.015**



Titer with 95% positive response based on the probit regression model

Probit regression is used to model dichotomous or binary outcome variables.
The spreadsheet requires 6 dilutions of a positive sample.

B. Measurement Uncertainty (Empirical Approaches)

Conformity assessment close to the LoD for a binary reporting qualitative examination based on "Uncertainty Approach" principles

Analyte:	HCV RNA	Units:	IU/mL
Notes:			

Serotec data base

no.	Sample ID	Date	Replicates Rep 1	Rep 2	Rep 3	Rep 4	Rep 5	Concentration
1	DU5748	June 26, 2017	1.5E+08	1.5E+08	1.5E+08	1.5E+08	1.5E+08	1.5E+08
2	OR3249	March 21, 2016	8.0E+07	8.1E+07	8.0E+07	7.8E+07	8.1E+07	8.0E+07
3	MK284	July 2, 2017	1.5E+08	1.6E+08	1.5E+08	1.5E+08	1.5E+08	1.5E+08
4	SJ1411	May 1, 2018	5.4E+07	5.2E+07	5.7E+07	5.7E+07	5.3E+07	5.4E+07
5	YG4110	May 28, 2016	7.0E+07	7.1E+07	7.1E+07	7.1E+07	6.9E+07	7.0E+07
6	LV1347	December 15, 2017	7.8E+07	7.7E+07	7.8E+07	7.6E+07	7.8E+07	7.7E+07
7	EC4883	July 5, 2017	3.1E+07	3.2E+07	3.2E+07	3.2E+07	3.2E+07	3.2E+07
8	VV4369	November 12, 2018	9.0E+07	9.0E+07	9.1E+07	8.9E+07	9.5E+07	9.1E+07
9	BC9192	January 4, 2017	1.5E+08	1.5E+08	1.5E+08	1.6E+08	1.5E+08	1.5E+08
10	BQ2020	January 7, 2014	2.0E+06	2.0E+06	2.1E+06	2.0E+06	2.0E+06	2.0E+06

Lowest mean concentration

	Sample ID	Date	Concentration
1st lowest:	GZ3826	July 28, 2017	1.8E+02
2nd lowest:	TX1179	November 17, 2017	9.5E+02
3rd lowest:	GL1723	May 29, 2018	1.0E+03
4th lowest:	JM8649	September 10, 2016	3.6E+05
5th lowest:	PM9734	September 19, 2018	4.1E+05

94	UX4247	November 4, 2016	7.4E+07	7.4E+07	7.5E+07	7.5E+07	7.3E+07	7.4E+07
95	HY6662	January 19, 2016	5.8E+07	5.6E+07	5.8E+07	5.8E+07	5.5E+07	5.7E+07
96	VJ9751	January 5, 2018	4.5E+07	4.6E+07	4.6E+07	4.5E+07	4.6E+07	4.6E+07
97	GZ3826	July 28, 2017	8.0E+01	3.0E+02	6.8E+01	3.4E+02	1.1E+02	1.8E+02
98	BJ715	May 2, 2019	6.1E+07	6.0E+07	6.0E+07	5.9E+07	6.2E+07	6.1E+07
99	ZG5141	November 9, 2016	1.3E+07	1.3E+07	1.3E+07	1.3E+07	1.3E+07	1.3E+07
100	LJ295	April 18, 2015	5.5E+05	5.5E+05	5.5E+05	5.6E+05	5.6E+05	5.5E+05

Test:	HCV RNA	Units:	IU/mL		
Technician:	PPereira	PPereira	PPereira	PPereira	PPereira
Date:	July 8, 2019	July 9, 2019	July 10, 2019	July 11, 2019	July 12, 2019
	Run 1	Run 2	Run 3	Run 4	Run 5
Rep 1	173.000	311.000	148.000	109.000	232.000
Rep 2	141.000	211.000	160.000	191.000	333.000
Rep 3	117.000	133.000	335.000	282.000	103.000
Rep 4	206.000	226.000	314.000	190.000	107.000
Rep 5	131.000	127.000	297.000	91.000	247.000
n =	5	5	5	5	5
df =	4	4	4	4	4
df =	20				
Sum =	768	1008	1254	863	1022
Mean =	153.6	201.6	250.8	172.6	204.4
Variance =	1282.8	5730.8	8007.7	5828.3	9720.8
SSW =	122281.6				

Source of variation	SS	DF	MS	n_0	V_B
Between-run =	27242.4	4	6810.6	5	139.304
Within-run =	122281.6	20	6114.08		
Total =	149524				

s_r =	78.19258277
s_I =	11.80271155
s_{Rw} =	**79.0783409**

The lowest value of 95% CI of the lowest concentration is 28 IU/mL. The value is higher than the lower limit of the linear range (8 IU/mL), so sample results up to 179 IU/mL are in the linear range with 95% confidence.

Test:	HCV RNA	Units:	IU/mL

Mean	Cert.value	*bias*	s_{bias}	n	U_{c_ref}	k	u_{c_ref}	u_{bias}
196.600	179.000	9.83%	18.8%	5	0	2	0.0%	**12.9%**

Run no.	Technician	Date	Replicate 1	Replicate 2	Replicate 3	Replicate 4	Replicate 5	Mean	*bias*	Relative *bias*
1	PPereira	July 8, 2019	173.000	141.000	117.000	206.000	131.000	153.600	-25.400	-14.2%
2	PPereira	July 9, 2019	311.000	211.000	133.000	226.000	127.000	201.600	22.600	12.6%
3	PPereira	July 10, 2019	148.000	160.000	335.000	314.000	297.000	250.800	71.800	40.1%
4	PPereira	July 11, 2019	109.000	191.000	282.000	190.000	91.000	172.600	-6.400	-3.6%
5	PPereira	July 12, 2019	232.000	333.000	103.000	107.000	247.000	204.400	25.400	14.2%

Test:	HCV RNA	Units:	IU/mL

s_{Rw}	u_{bias}	u	k	U
40.2%	12.9%	42.2%	2.00	**84.5%**

C. Total Analytical Error

The calculation of the standard deviation and the bias is the same as for the measurement uncertainty. What differs is the mathematical model for combining the total analytical error.

Single-laboratory validation approach using short-term data

Test:	anti-HCV	Units:	S/CO

s_{Rw}	*bias*	*k*	**TAE**
6.3%	1.2%	2.00	**13.9%**

Intralaboratory approach using short-term data

Single-laboratory validation approach using long-term data

Test:	anti-HCV	Units:	S/CO

s_{Rw}	*bias*	*k*	**TAE**
9.0%	1.2%	2.00	**19.3%**

Intralaboratory approach using long-term data

Interlaboratory comparisons

Using intercomparison data

Test:	anti-HCV	Units:	S/CO

Reference result	Lab result	Deviation	s_R		No. of ref labs	Excluded
0.980	0.992	1.2%	0.180	18.4%	2	0

s_{Rw}	*k*	**TAE**
18.4%	2.00	**38.0%**

External quality assessment / proficiency testing

Test:	anti-HCV	Units:	S/CO

s_{Rw}	*bias*	*k*	**TAE**
9.0%	-2.4%	2.00	**20.4%**

Using EQA / PT data for bias estimate

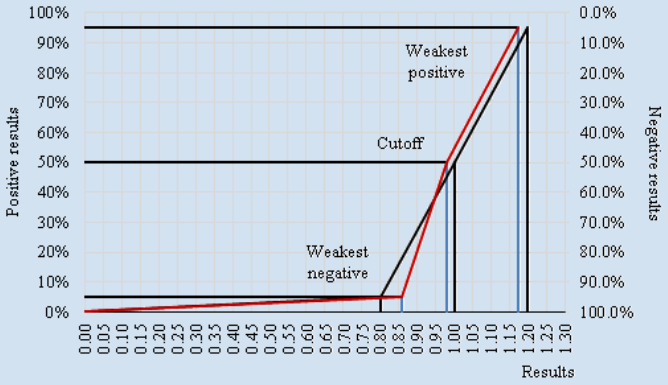
Combination with within-laboratory reproducibility standard deviation

D. C_5-C_{95} Interval of the Cutoff Value

Test:	Immunoassay	Units:	S/CO
Technician:	PPereira		
Date:	March 24, 2018		
Cutoff:	1	Range:	20.0%

	C_{50}+%		C_{50}		C_{50}-%	
no.	Results	Binary	Results	Binary	Results	Binary
1	1.06	Positive	0.98	Negative	0.88	Negative
2	1.26	Positive	0.99	Negative	0.83	Negative
3			0.95	Negative	0.91	Negative
4	1.23	Positive	0.97	Negative	0.86	Negative
5	1.11	Positive	0.98	Negative	0.79	Negative
6	1.15	Positive	0.97	Negative	0.88	Negative
7	1.21	Positive	0.88	Negative	0.86	Negative
8	1.13	Positive	1.02	Positive	0.87	Negative
9	1.21	Positive	0.98	Negative	0.87	Negative
10	1.17	Positive	1.02	Positive	0.86	Negative
11	1.19	Positive	0.97	Negative	0.86	Negative
12	1.15	Positive	1.06	Positive	0.91	Negative
13	1.14	Positive	0.97	Negative	0.80	Negative
14	1.10	Positive	1.04	Positive	0.90	Negative
15	1.23	Positive	0.95	Negative	0.79	Negative
16	1.25	Positive	1.04	Positive	0.89	Negative
17	1.31	Positive	0.93	Negative	0.87	Negative
18	1.08	Positive	0.93	Negative	0.86	Negative
19	1.14	Positive	0.99	Negative	0.86	Negative
20	1.20	Positive	0.95	Negative	0.83	Negative

n of 40 is required for the three sample



Graphical illustration of the bias at the three levels

Percentage of binary results					
C_{50}+%		C_{50}		C_{50}-%	
Positive	Negative	Positive	Negative	Positive	Negative
100.0%	0.0%	37.5%	62.5%	0.0%	100.0%

	C_{50}+%	C_{50}	C_{50}-%
n	40	40	40
Average	1.175	0.979	0.858
SD	0.062	0.049	0.040
1.96*SD	0.121	0.096	0.078
CV	5.3%	5.0%	4.6%

Compliance assessment:
Was the prepared C50 incorrect?
Correct
Does the concentration range bound (contain) the C5-C95 interval for the candidate method?
Bounds the C5-C95 interval

Compliance assessment

E. Condition Accuracy

Evaluation of the condition accuracy of a single test

Test:	anti-HCV	Units:	Binary
Technician:	PPereira		
Date:	July 17, 2019		

Condition sensitivity

2X2 contingency table

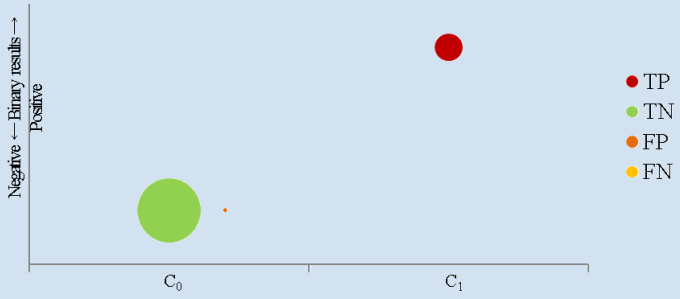
C1Sample	Result	C0Sample	Result
sa1	+	sb1	-
sa2	+	sb2	-
sa3	+	sb3	-
sa4	+	sb4	-
sa5	+	sb5	-
sa6	+	sb6	-
sa7	+	sb7	-
sa8	+	sb8	-
sa9	+	sb9	-
sa10	+	sb10	-
sa11	+	sb11	-
sa12	+	sb12	-
sa13	+	sb13	-
sa14	+	sb14	-
sa15	+	sb15	-
sa16	+	sb16	-
sa17	+	sb17	+
sa18	+	sb18	-
sa19	+	sb19	-
sa20	+	sb20	-
sa21	+	sb21	-
sa22	+	sb22	-
sa23	+	sb23	-
sa24	+	sb24	-
sa25	+	sb25	-
sa31	+	sb31	-
sa32	+	sb32	-

	Condition accuracy		
Candidate assay	Positive	Negative	Total
Positive	49	1	50
Negative	0	256	256
Total	49	257	306

Sensitiviy=	1	**Specificity=**	0.996108949
A1=	101.8416	B1=	515.8416
A2=	3.8416	B2=	5.483110351
A3=	185.6832	B3=	521.6832
LL=	0.927299703	**LL=**	0.978291978
HL=	1	**HL=**	0.999312821
PPV=	0.98	NPV=	0.98
Prevalence=	0.160130719	Efficiency=	0.996732026

Condition specificity

No. of patients *n* =	250000
Population *N*	10000000
PDP=	0.03
PPV=	0.868243243
NPV=	1
E=	0.024470215
FDR=	0.131756757
FOR=	0



Bubble plots graphic

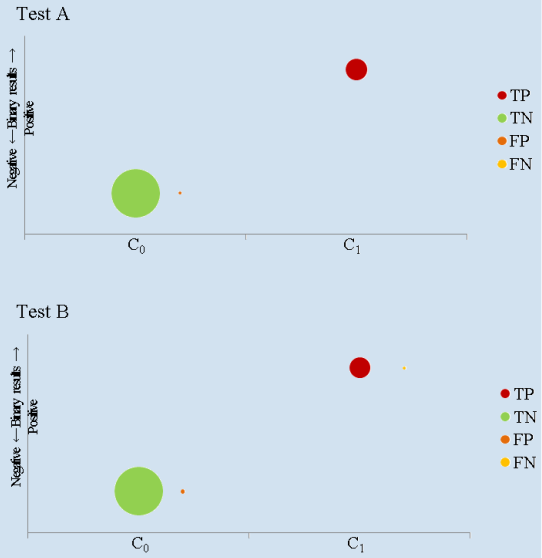
E. Condition Accuracy

Comparison of the condition accuracy of two tests

Test:	anti-HCV
Technician:	PPereira
Date:	July 18, 2019

Units:	Binary

no.	C1Sample	C1	Test A	Test B	C0Sample	C0	Test A	Test B
1	sa1	+	+	+	sb1	-	-	-
2	sa2	+	+	+	sb2	-	-	-
3	sa3	+	+	+	sb3	-	-	-
4	sa4	+	+	+	sb4	-	-	-
5	sa5	+	+	+	sb5	-	-	-
6	sa6	+	+	+	sb6	-	-	-
7	sa7	+	+	+	sb7	-	-	-
8	sa8	+	+	+	sb8	-	-	-
9	sa9	+	+	+	sb9	-	-	+
10	sa10	+	+	+	sb10	-	-	-
11	sa11	+	+	+	sb11	-	-	-
12	sa12	+	+	+	sb12	-	-	-
13	sa13	+	+	-	sb13	-	-	-
14	sa14	+	+	+	sb14	-	-	-
15	sa15	+	+	+	sb15	-	-	-
16	sa16	+	+	+	sb16	-	-	-
17	sa17	+	+	+	sb17	-	+	+
18	sa18	+	+	+	sb18	-	-	-
19	sa19	+	+	+	sb19	-	-	-
20	sa20	+	+	+	sb20	-	-	-
21	sa21	+	+	+	sb21	-	-	-
22	sa22	+	+	+	sb22	-	-	-
23	sa23	+	+	+	sb23	-	-	-
24	sa24	+	+	+	sb24	-	-	-
25	sa25	+	+	+	sb25	-	-	-
26	sa26	+	+	+	sb26	-	-	-
27	sa27	+	+	+	sb27	-	-	-
28	sa28	+	+	+	sb28	-	-	-
29	sa29	+	+	+	sb29	-	-	-
30	sa30	+	+	+	sb30	-	-	-



Test A	Condition accuracy		
Candidate assay	Positive	Negative	Total
Positive	49	1	50
Negative	0	256	256
Total	49	257	306

Test B	Condition accuracy		
Candidate assay	Positive	Negative	Total
Positive	48	2	50
Negative	1	255	256
Total	49	257	306

Sensitiviy=	1	**Specificity=**	0.996108949
A1=	101.8416	B1=	515.8416
A2=	3.8416	B2=	5.483110351
A3=	105.6832	B3=	521.6832
LL=	0.927299703	**LL=**	0.978291978
HL=	1	**HL=**	0.999312821
PPV=	0.98	NPV=	0.98
Prevalence=	0.160130719	Efficiency=	0.996732026

Sensitiviy=	0.979592	**Specificity=**	0.992217899
A1=	99.8416	B1=	513.8416
A2=	5.459917	B2=	6.72692536
A3=	105.6832	B3=	521.6832
LL=	0.893062	**LL=**	0.972073999
HL=	0.996388	**HL=**	0.997863311
PPV=	0.96	NPV=	0.96
Prevalence=	0.160131	Efficiency=	0.990196078

SensitivityTest1=	1	SpecificityTest1=	0.99
SensitivityTest2=	0.979591837	SpecificityTest2=	0.98
SeTest1-SeTest2=	0.020408163	**SpTest1-SpTest2=**	0.01
A1=	0	B1=	19404
A2=	0	B2=	98
A3=	0	B3=	48
A4=	0	B4=	0.344584394
A5=	0.005567458	B5=	0.000123397
A6=	0.007487358	B6=	0.000371564
HL=	-0.05420724	**HL=**	-0.001108408
LL=	0.106937685	**LL=**	0.029275985

Differences in condition sensitivity and specificity

F. Condition Accuracy by Delta-Value

Test 1

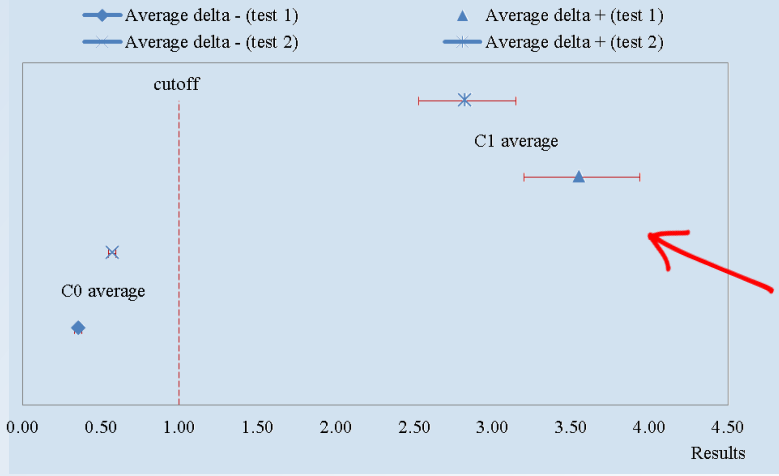
Test:	anti-HCV	Units:	S/CO
Technician:	PPereira		
Date:	September 12, 2019		

	Samples		C1		C0		
no.	C1	C0	Results	log10	Results	log10	
1	sa1	sb1	5.42	0.734	1.20	0.079	Log10aver= 0.550
2	sa2	sb2	3.98	0.600	0.37	-0.432	Log10s= 0.161
3	sa3	sb3	3.29	0.517	0.23	-0.638	**d+= 3.416**
4	sa4	sb4	3.45	0.538	0.23	-0.638	Log10_95%CI= 0.045
5	sa5	sb5	3.85	0.585	0.64	-0.194	Log10HL= 0.595
6	sa6	sb6	2.48	0.394	0.69	-0.161	Log10LL= 0.505
7	sa7	sb7	2.10	0.322	0.13	-0.886	Average= 3.549
8	sa8	sb8	3.80	0.580	0.48	-0.319	HL= 3.938
9	sa9	sb9	4.00	0.602	0.38	-0.420	LL= 3.199
10	sa10	sb10	3.86	0.587	0.19	-0.721	Log10aver= -0.450
11	sa11	sb11	2.71	0.433	0.16	-0.796	Log10s= 0.232
12	sa12	sb12	4.76	0.678	0.27	-0.569	**d-= -1.934**
13	sa13	sb13	2.60	0.415	0.42	-0.377	Log10_95%CI= 0.028
14	sa14	sb14	4.52	0.655	0.69	-0.161	Log10HL= -0.421
15	sa15	sb15	5.45	0.736	0.11	-0.959	Log10LL= -0.478
16	sa16	sb16	3.93	0.594	0.45	-0.347	Average= 0.355
17	sa17	sb17	5.60	0.748	0.35	-0.456	HL= 0.379
18	sa18	sb18	3.64	0.561	0.21	-0.678	LL= 0.333
19	sa19	sb19	2.59	0.413	0.26	-0.585	
20	sa20	sb20	5.33	0.727	0.30	-0.523	

Test 2

Test:	anti-HCV	Units:	S/CO
Technician:	PPereira		
Date:	September 12, 2019		

C1		C0		
Results	log10	Results	log10	
1.2	0.079	4	0.602	Log10aver= 0.451
1.64	0.215	0.94	-0.027	Log10s= 0.170
3.83	0.583	0.62	-0.208	**d+= 2.658**
4.39	0.642	0.76	-0.119	Log10_95%CI= 0.047
4.76	0.678	0.42	-0.377	Log10HL= 0.498
2.83	0.452	0.38	-0.420	Log10LL= 0.403
1.73	0.238	0.89	-0.051	Average= 2.822
4.54	0.657	0.85	-0.071	HL= 3.148
4.33	0.636	0.88	-0.056	LL= 2.530
2.24	0.350	0.62	-0.208	Log10aver= -0.243
2.01	0.303	0.9	-0.046	Log10s= 0.150
4.66	0.668	0.55	-0.260	**d-= -1.616**
2.78	0.444	0.58	-0.237	Log10_95%CI= 0.018
3.14	0.497	0.77	-0.114	Log10HL= -0.224
2.55	0.407	0.33	-0.481	Log10LL= -0.261
4.85	0.686	0.5	-0.301	Average= 0.572
1.8	0.255	0.9	-0.046	HL= 0.596
2.05	0.312	0.67	-0.174	LL= 0.548
2.01	0.303	0.59	-0.229	
3.33	0.522	0.35	-0.456	



Graphical illustration of the averages and 95% CI for the weakest results

G. Seroconversion Window Period

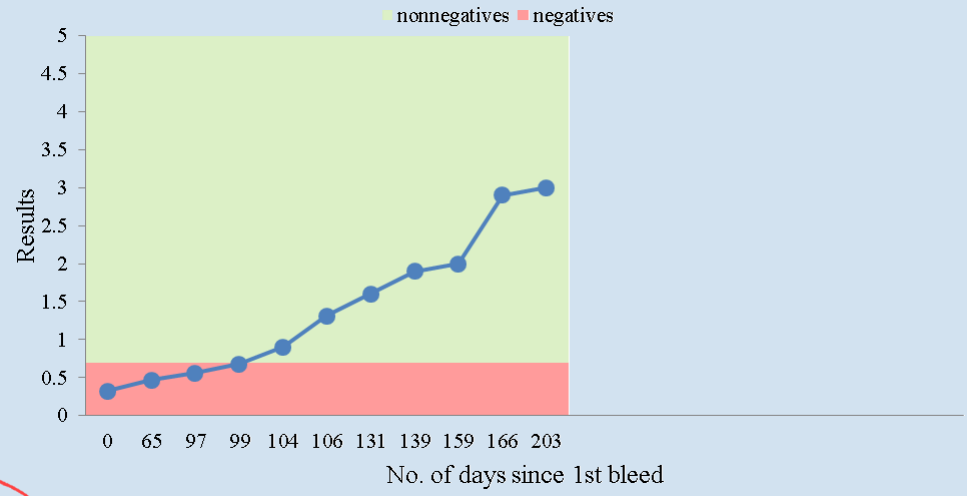
Test:	anti-HCV
Technician:	PPereira
Date:	March 12, 2019

Units:	S/CO

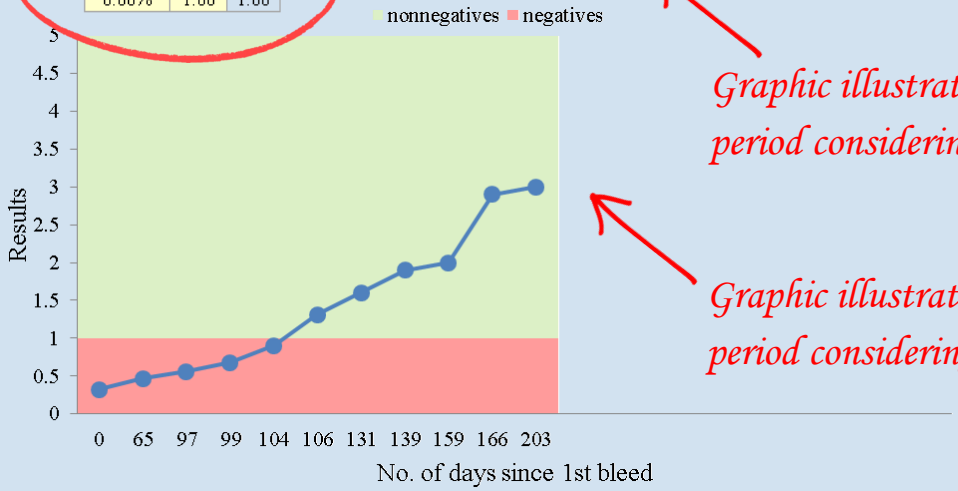
gray-zone	cutoff	DL
30.00%	1.00	0.70

no.	Sample ID	Sample collection dates	No. of days since 1st collect	Results	Trinary result	Window period (No. of days)
1	PHV901-01	9/23/1993	0	0.32	Negative	0
2	PHV901-02	11/27/1993	65	0.47	Negative	65
3	PHV901-03	12/29/1993	97	0.56	Negative	97
4	PHV901-04	12/31/1993	99	0.68	Negative	99
5	PHV901-05	1/5/1994	104	0.90	Indeterminate	104
6	PHV901-06	1/7/1994	106	1.31	Positive	
7	PHV901-07	2/1/1994	131	1.60	Positive	
8	PHV901-08	2/9/1994	139	1.90	Positive	
9	PHV901-09	3/1/1994	159	2.00	Positive	
10	PHV901-10	3/8/1994	166	2.90	Positive	
11	PHV901-11	4/14/1994	203	3.00	Positive	
12						
13						
14						
15						
16						
17						
18						
19						
20						

Number of days after day-zero



gray-zone	cutoff	DL
0.00%	1.00	1.00



Graphic illustration of the seroconversion period considering the "gray zone"

Graphic illustration of the seroconversion period considering the cutoff point

H. Agreement of Binary Results

Test:	anti-HCV	Units:	Binary
Technician:	PPereira		
Date:	July 17, 2019		

no.	C1Sample	Result	C0Sample	Result
1	sa1	+	sb1	-
2	sa2	+	sb2	-
3	sa3	+	sb3	-
4	sa4	+	sb4	-
5	sa5	+	sb5	-
6	sa6	+	sb6	-
7	sa7	+	sb7	-
8	sa8	-	sb8	-
9	sa9	+	sb9	-
10	sa10	+	sb10	-
11	sa11	+	sb11	-
12	sa12	+	sb12	-
13	sa13	+	sb13	-
14	sa14	+	sb14	-
15	sa15	+	sb15	-
16	sa16	+	sb16	-
17	sa17	+	sb17	-
18	sa18	+	sb18	-
19	sa19	+	sb19	-
20	sa20	+	sb20	+

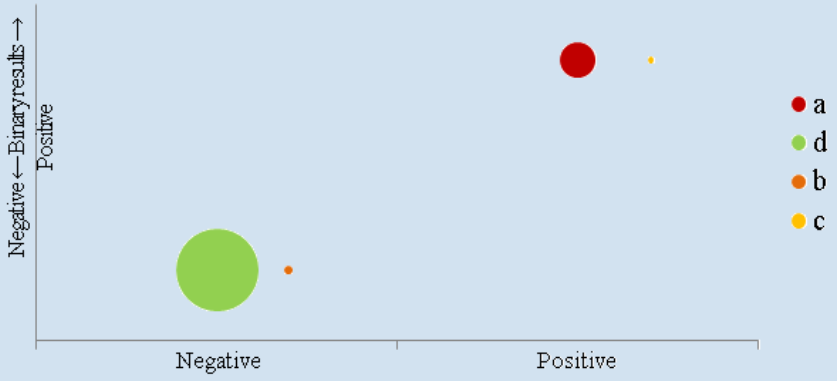
2X2 contingency table

	Comparative test results		
Candidate test	Positive	Negative	Total
Positive	37	2	39
Negative	1	196	197
Total	38	198	236

PA(a)=	0.974	PNA(d)=	0.990
A1=	77.842	B1=	395.842
A2=	5.452	B2=	6.722
A3=	83.683	B3=	403.683
LL=	0.865	LL=	0.964
HL=	0.995	HL=	0.997
OA =	0.987		

Negative agreement

Positive agreement



Bubble plots graphic

I. Agreement of Binary Results

Area under the curve

Youden index

Test:	anti-HCV
Technician:	PPereira
Date:	September 26, 2019

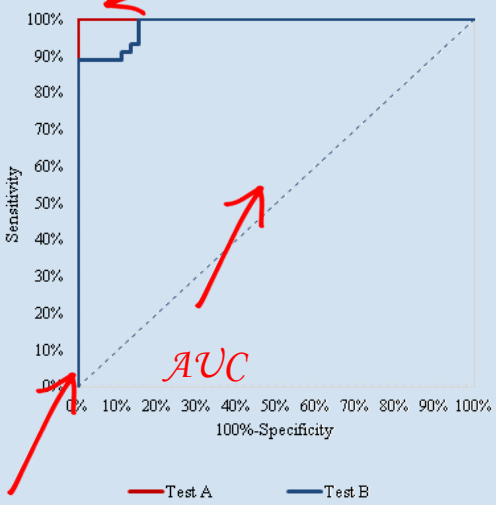
Units: A

AUC = 1.000 | LL = 0.999 | HL = 1.000

Youden Index = 1.000

	Samples		Results									95% CI			95% CI				
no.	D_1	D_0	D_1	D_0	n_p	p_n	TP	TN	FN	FP	N	se	LL	HL	sp	LL	HL	1-sp	TPR	FPR
4	sa4	sab4	1620	90	0.003	10	45	0	1	44	90	100.0%	92.1%	100.0%	2.2%	0.4%	11.6%	97.8%	100.0%	97.8%
5	sa5	sab5	2304	108	0.004	10	45	0	1	44	90	100.0%	92.1%	100.0%	2.2%	0.4%	11.6%	97.8%	100.0%	97.8%
6	sa6	sab6	1269	16	0.005	10	45	0	1	44	90	100.0%	92.1%	100.0%	2.2%	0.4%	11.6%	97.8%	100.0%	97.8%
7	sa7	sab7	1317	36	0.006	11	45	0	1	44	90	100.0%	92.1%	100.0%	2.2%	0.4%	11.6%	97.8%	100.0%	97.8%
8	sa8	sab8	1372	29	0.007	11	45	0	1	44	90	100.0%	92.1%	100.0%	2.2%	0.4%	11.6%	97.8%	100.0%	97.8%
9	sa9	sab9	2145	66	0.008	11	45	0	1	44	90	100.0%	92.1%	100.0%	2.2%	0.4%	11.6%	97.8%	100.0%	97.8%
10	sa10	sab10	2309	11	0.009	11	45	0	1	44	90	100.0%	92.1%	100.0%	2.2%	0.4%	11.6%	97.8%	100.0%	97.8%
11	sa11	sab11	1175	37	0.010	11	45	0	1	44	90	100.0%	92.1%	100.0%	2.2%	0.4%	11.6%	97.8%	100.0%	97.8%
12	sa12	sab12	1044	32	0.011	11	45	0	1	44	90	100.0%	92.1%	100.0%	2.2%	0.4%	11.6%	97.8%	100.0%	97.8%
13	sa13	sab13	2126	138	0.012	11	45	0	2	43	90	100.0%	92.1%	100.0%	4.4%	1.2%	14.8%	95.6%	100.0%	95.6%
14	sa14	sab14	1918	45	0.013	12	45	0	2	43	90	100.0%	92.1%	100.0%	4.4%	1.2%	14.8%	95.6%	100.0%	95.6%
15	sa15	sab15	1747	70	0.014	12	45	0	2	43	90	100.0%	92.1%	100.0%	4.4%	1.2%	14.8%	95.6%	100.0%	95.6%
16	sa16	sab16	1593	67	0.015	13	45	0	2	43	90	100.0%	92.1%	100.0%	4.4%	1.2%	14.8%	95.6%	100.0%	95.6%
17	sa17	sab17	1163	44	0.016	13	45	0	2	43	90	100.0%	92.1%	100.0%	4.4%	1.2%	14.8%	95.6%	100.0%	95.6%
18	sa18	sab18	1471	10	0.017	14	45	0	2	43	90	100.0%	92.1%	100.0%	4.4%	1.2%	14.8%	95.6%	100.0%	95.6%
19	sa19	sab19	2147	106	0.018	14	45	0	2	43	90	100.0%	92.1%	100.0%	4.4%	1.2%	14.8%	95.6%	100.0%	95.6%
20	sa20	sab20	1685	68	0.019	14	45	0	2	43	90	100.0%	92.1%	100.0%	4.4%	1.2%	14.8%	95.6%	100.0%	95.6%
494					0.493	145	45	0	44	1	90	100.0%	92.1%	100.0%	97.8%	88.4%	99.6%	2.2%	100.0%	2.2%
495					0.494	145	45	0	44	1	90	100.0%	92.1%	100.0%	97.8%	88.4%	99.6%	2.2%	100.0%	2.2%
496					0.495	194	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
497					0.496	274	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
498					0.497	354	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
499					0.498	434	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
500					0.499	514	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
501					0.500	595	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
502					0.501	675	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
503					0.502	755	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
504					0.503	835	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
505					0.504	915	45	0	45	0	90	100.0%	92.1%	100.0%	100.0%	92.1%	100.0%	0.0%	100.0%	0.0%
506					0.505	995	45	1	45	0	91	97.8%	88.7%	99.6%	100.0%	92.1%	100.0%	0.0%	97.8%	0.0%
507					0.506	1046	44	1	45	0	90	97.8%	88.4%	99.6%	100.0%	92.1%	100.0%	0.0%	97.8%	0.0%

Tests' comparison

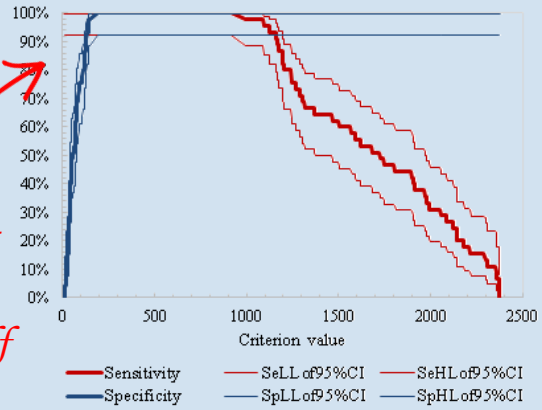


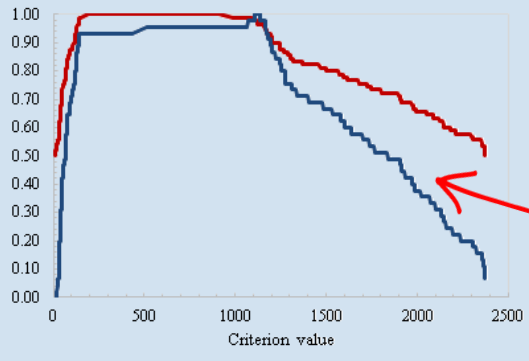
AUC

ROC

Identification of cutoff point

Graphical illustration of the condition sensitivity and specificity for hypothetical cutoff





Graphic illustration of efficiency and Youden index for hypothetical cutoff points

J. IQC / EQA: Sigma metrics Based on DPMO

DPMO-derived Sigma Metrics: Sigma based on the number of defects per million opportunities

Test: anti-HCV | Legend ($sigma_{DPMO}$ and $simg_{DPMO}$-1.5): ≥4 | [3, 4[| <3

Notes: Defects = human samples reject due rejected analytical runs | Opportunities = human samples tested | Defect opport. = three lab processes

no.	Technician	Date From	Date To	no. of days	Defects	Opportunities	Defect opport. per unit
1	PPereira	5/12/2018	5/12/2019	365	225	81450	3
2	PPereira	7/14/2018	1/30/2019	200	30	12670	3
3	PPereira	5/12/2018	5/12/2019	365	10	92466	3
4	PPereira	5/12/2018	5/12/2019	365	1	105332	3
5	PPereira	5/12/2018	5/12/2019	365	225	81450	3
6	PPereira	5/12/2018	5/12/2019	365	200	92458	3
7	PPereira	5/12/2018	5/12/2019	365	190	87212	3
8	PPereira	5/12/2018	5/12/2019	365	980	22568	3
9	PPereira	5/12/2018	5/12/2019	365	4470	21679	3

Compliance assessment

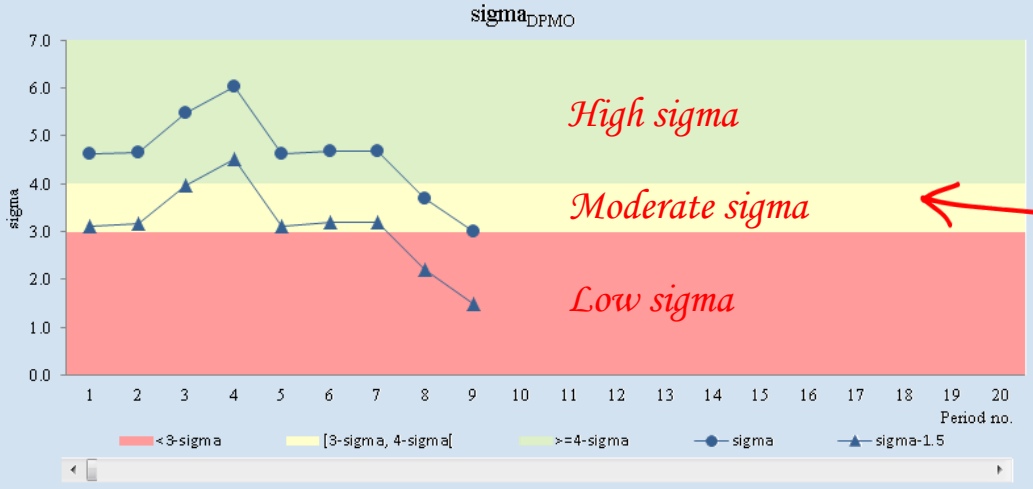
Possibility of calculation in various scenarios, being able to be representative of the entire laboratory process and not only of the analytical process

DPMO	$sigma_{DPMO}$	Performance
920.8104	4.62	Satisfactory process - meets specification limits
789.266	4.66	Satisfactory process - meets specification limits
36.6493	5.47	Satisfactory process - meets specification limits
3.1646	6.02	Satisfactory process - meets specification limits
920.8104	4.62	Satisfactory process - meets specification limits
721.0482	4.69	Satisfactory process - meets specification limits
726.1998	4.69	Satisfactory process - meets specification limits
14474.7726	3.69	Capable process,but marginally - process will not tolerate a significant shift
68730.1075	2.99	The process is out of specificaiton or about to happen

Short-term sigma metrics

Long-term sigma metrics ("real sigma")

$sigma_{DPMO}$-1.5	Performance
3.12	Capable process,but marginally - process will not tolerate a significant shift
3.16	Capable process,but marginally - process will not tolerate a significant shift
3.97	Capable process,but marginally - process will not tolerate a significant shift
4.52	Satisfactory process - meets specification limits
3.12	Capable process,but marginally - process will not tolerate a significant shift
3.19	Capable process,but marginally - process will not tolerate a significant shift
3.19	Capable process,but marginally - process will not tolerate a significant shift
2.19	The process is out of specificaiton or about to happen
1.49	The process is out of specificaiton or about to happen



Graphical illustration of sigma metrics

K. IQC / EQA: Sigma Metrics Based on Systematic Error

SE-derived Sigma Metrics: Sigma based on the critical systematic error

Allowable loss of sensitivity

Test: anti-HCV | Units: S/CO | Sample concentration limits: Min. conc. = 1.50 | Max. conc. = 3.00

no.	Technician	Date	Sample	s_{RW}	Average	cutoff	"gray zone"	Decision Limit	ALS Absolute	ALS Percentage
1	PPereira	5/12/2015	sample1	0.05	2.30	1.00	10.0%	0.90	1.40	140.0%
2	PPereira	5/13/2015	sample2	0.05	2.10	1.00	10.0%	0.90	1.20	120.0%
3	PPereira	5/14/2015	sample3	0.05	2.50	1.00	10.0%	0.90	1.60	160.0%
4	PPereira	5/15/2015	sample4	0.05	1.70	1.00	10.0%	0.90	0.80	80.0%
5	PPereira	5/16/2015	sample5	0.05	2.60	1.00	10.0%	0.90	1.70	170.0%
6	PPereira	5/17/2015	sample6	0.05	2.30	1.00	10.0%	0.90	1.40	140.0%
7	PPereira	5/18/2015	sample7	0.05	2.60	1.00	10.0%	0.90	1.70	170.0%
8	PPereira	5/12/2015	sample8	0.05	1.10	1.00	10.0%	0.90	0.20	20.0%
9	PPereira	5/13/2015	sample9	0.05	1.70	1.00	10.0%	0.90	0.80	80.0%
10	PPereira	5/14/2015	sample10	0.05	2.80	1.00	10.0%	0.90	1.90	190.0%
11	PPereira	5/15/2015	sample11	0.05	1.60	1.00	10.0%	0.90	0.70	70.0%
12	PPereira	5/16/2015	sample12	0.05	2.85	1.00	10.0%	0.90	1.95	195.0%
13	PPereira	5/17/2015	sample13	0.05	2.90	1.00	10.0%	0.90	2.00	200.0%
14	PPereira	5/18/2015	sample14	0.05	1.70	1.00	10.0%	0.90	0.80	80.0%
15	PPereira	5/19/2015	sample15	0.05	1.50	1.00	10.0%	0.90	0.60	60.0%
16	PPereira	5/20/2015	sample16	0.05	2.10	1.00	10.0%	0.90	1.20	120.0%
17	PPereira	5/21/2015	sample17	0.05	1.50	1.00	10.0%	0.90	0.60	60.0%
18	PPereira	5/22/2015	sample18	0.05	2.00	1.00	10.0%	0.90	1.10	110.0%
19	PPereira	5/23/2015	sample19	0.05	2.70	1.00	10.0%	0.90	1.80	180.0%
20	PPereira	5/24/2015	sample20	0.05	2.40	1.00	10.0%	0.90	1.50	150.0%

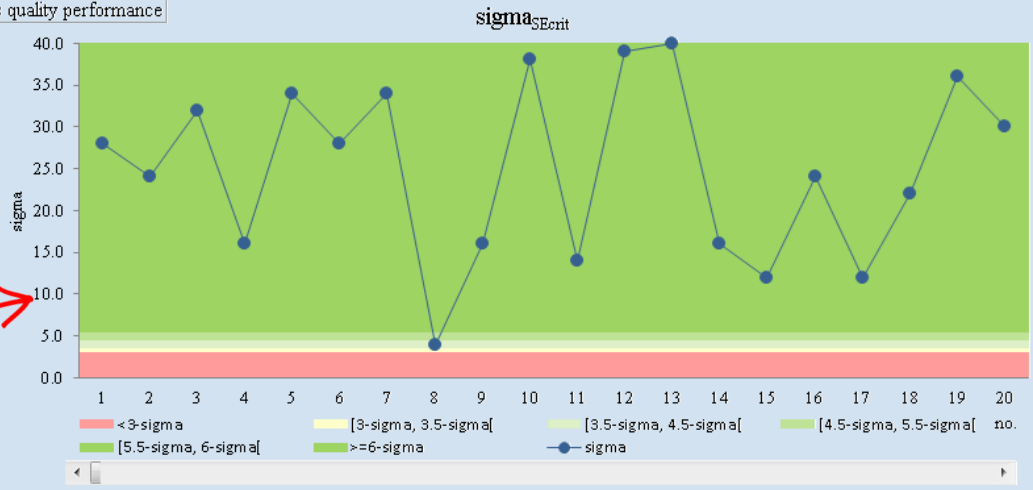
SE_{crit}	$sigma_{crit}$	Performance
26.35	28	World-class quality performance
22.35	24	World-class quality performance
30.35	32	World-class quality performance
14.35	16	World-class quality performance
32.35	34	World-class quality performance
26.35	28	World-class quality performance
32.35	34	World-class quality performance
2.35	4	Moderate to low performance
14.35	16	World-class quality performance
36.35	38	World-class quality performance
12.35	14	World-class quality performance
37.35	39	World-class quality performance
38.35	40	World-class quality performance
14.35	16	World-class quality performance
10.35	12	World-class quality performance
22.35	24	World-class quality performance
10.35	12	World-class quality performance
20.35	22	World-class quality performance
34.35	36	World-class quality performance
28.35	30	World-class quality performance

SE_{crit} and SE-derived Sigma Metrics

Legend ($sigma_{crit}$):		
≥6	[5.5, 6[	[4.5, 5.5[
[3.5, 4.5[	[3.5 - 3]	<3

Compliance assessment

Graphical illustration of sigma metrics



L. IQC / EQA: Internal Quality Control (Numerical Data)

Qualitative results classified in an ordinal scale

Accepted results and z value

Test:	Immunoassay			Units:	S/CO						
Stats:	Moving average = 2.42			Mean of s =	0.15	CV = 6.0%		n = 292		Startup of	30

no.	Date	Time	Technician	Reag. kit batch	Raw data	Calc.value	z-value	Material	Lot batch	Mean	s
1	5/12/2015	2:52:24 PM	PPereira	12345ABC67	2.44	2.44	0.00	M1	M1234	2.44	0.00
2	5/13/2015	3:52:24 PM	PPereira	12345ABC67	2.47	2.47	0.00	M1	M1234	2.47	0.00
3	5/14/2015	4:52:24 PM	PPereira	12345ABC67	2.20	2.20	-12.02	M1	M1234	2.20	0.15
4	5/15/2015	5:52:24 PM	PPereira	12345ABC67	2.64	2.64	1.82	M1	M1234	2.64	0.18
5	5/16/2015	6:52:24 PM	PPereira	12345ABC67	2.50	2.50	0.34	M1	M1234	2.50	0.16
6	5/17/2015	7:52:24 PM	PPereira	12345ABC67	2.44	2.44	-0.06	M1	M1234	2.44	0.14
7	5/18/2015	8:52:24 PM	PPereira	12345ABC67	2.35	2.35	-0.69	M1	M1234	2.35	0.14
8	5/19/2015	9:52:24 PM	PPereira	12345ABC67	3.71			M1	M1234		
9	5/20/2015	10:52:24 PM	PPereira	12345ABC67	2.35	2.35	-0.62	M1	M1234	2.35	0.13
10	5/21/2015	11:52:24 PM	PPereira	12345ABC67	2.44	2.44	0.13	M1	M1234	2.44	0.12
11	5/22/2015	12:52:24 AM	PPereira	12345ABC67	2.61	2.61	1.53	M1	M1234	2.44	0.13
12	5/23/2015	1:52:24 AM	PPereira	12345ABC67	2.41	2.41	-0.27	M1	M1234	2.44	0.12
13	5/24/2015	2:52:24 AM	PPereira	12345ABC67	2.47	2.47	0.24	M1	M1234	2.44	0.12
14	5/25/2015	3:52:24 AM	PPereira	12345ABC67	2.25	2.25	-1.66	M1	M1234	2.43	0.12
15	5/26/2015	4:52:24 AM	PPereira	12345ABC67	2.26	2.26	-1.36	M1	M1234	2.42	0.13
16	5/27/2015	5:52:24 AM	PPereira	12345ABC67	2.35	2.35	-0.52	M1	M1234	2.41	0.12
17	5/28/2015	6:52:24 AM	PPereira	12345ABC67	2.41	2.41	-0.02	M1	M1234	2.41	0.12
18	5/29/2015	7:52:24 AM	PPereira	12345ABC67	2.44	2.44	0.24	M1	M1234	2.41	0.12
19	5/30/2015	8:52:24 AM	PPereira	12345ABC67	2.40	2.40	-0.12	M1	M1234	2.41	0.11
20	5/31/2015	9:52:24 AM	PPereira	12345ABC67	2.33	2.33	-0.74	M1	M1234	2.41	0.11
21	6/1/2015	10:52:24 AM	PPereira	12345ABC67	2.56	2.56	1.37	M1	M1234	2.42	0.11
22	6/2/2015	11:52:24 AM	PPereira	12345ABC67	2.44	2.44	0.21	M1	M1234	2.42	0.11
23	6/3/2015	12:52:24 PM	PPereira	12345ABC67	2.31	2.31	-0.97	M1	M1234	2.41	0.11
24	6/4/2015	1:52:24 PM	PPereira	12345ABC67	2.53	2.53	1.07	M1	M1234	2.42	0.11
25	6/5/2015	2:52:24 PM	PPereira	12345ABC67	2.58	2.58	1.47	M1	M1234	2.42	0.11
26	6/6/2015	3:52:24 PM	PPereira	12345ABC67	2.28	2.28	-1.28	M1	M1234	2.42	0.11
27	6/7/2015	4:52:24 PM	PPereira	12345ABC67	2.39	2.39	-0.25	M1	M1234	2.42	0.11
28	6/8/2015	5:52:24 PM	PPereira	12345ABC67	2.50	2.50	0.74	M1	M1234	2.42	0.11
29	6/9/2015	6:52:24 PM	PPereira	12345ABC67	2.35	2.35	-0.63	M1	M1234	2.42	0.11
30	6/10/2015	7:52:24 PM	PPereira	12345ABC67	2.43	2.43	0.11	M1	M1234	2.42	0.11
31	6/11/2015	8:52:24 PM	PPereira	12345ABC67	2.60	2.60	1.69	M1	M1234	2.42	0.11
32	6/12/2015	9:52:24 PM	PPereira	12345ABC67	2.40	2.40	-0.22	M1	M1234	2.42	0.11
33	6/13/2015	10:52:24 PM	PPereira	12345ABC67	2.45	2.45	0.24	M1	M1234	2.42	0.11
34	6/14/2015	11:52:24 PM	PPereira	12345ABC67	2.60	2.60	1.63	M1	M1234	2.43	0.11
35	6/15/2015	12:52:24 AM	PPereira	12345ABC67	2.66	2.66	2.09	M1	M1234	2.44	0.12
36	6/16/2015	1:52:24 AM	PPereira	12345ABC67	2.62	2.62	1.59	M1	M1234	2.44	0.12
37	6/17/2015	2:52:24 AM	PPereira	12345ABC67	2.66	2.66	1.85	M1	M1234	2.45	0.12

Reagent traceability

QC material traceability

Rules' selection

Rejection or warning notice

Rule or multi-rules selection (mark "x")

x	x											Rule		
1:3s	1:2s	2:2s	R:4s	4:1s	6x	8x	9x	10x	12x	2of3:2	7:T	violation	Excl	Actions
1:3s	1:2s											Reject		
													i	CAPA 19/2019

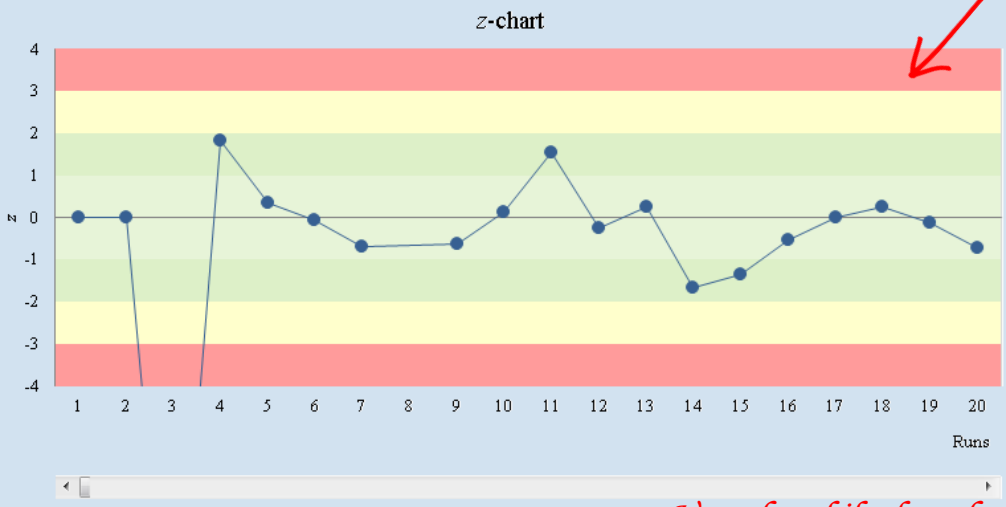
Possibility of rejection of results

Actions taken

L. IQC / EQA: Internal Quality Control (Numerical Data)

Internal quality control

Z-chart for a more natural interpretation of the rejections and warnings



Use of mobile data for a more representative illustration of the stability of the test results and greater robustness in the application of the rules

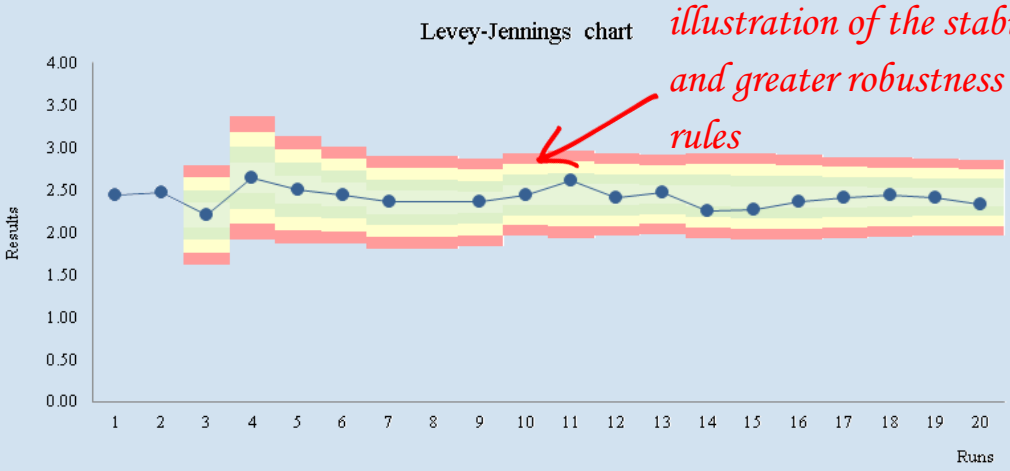
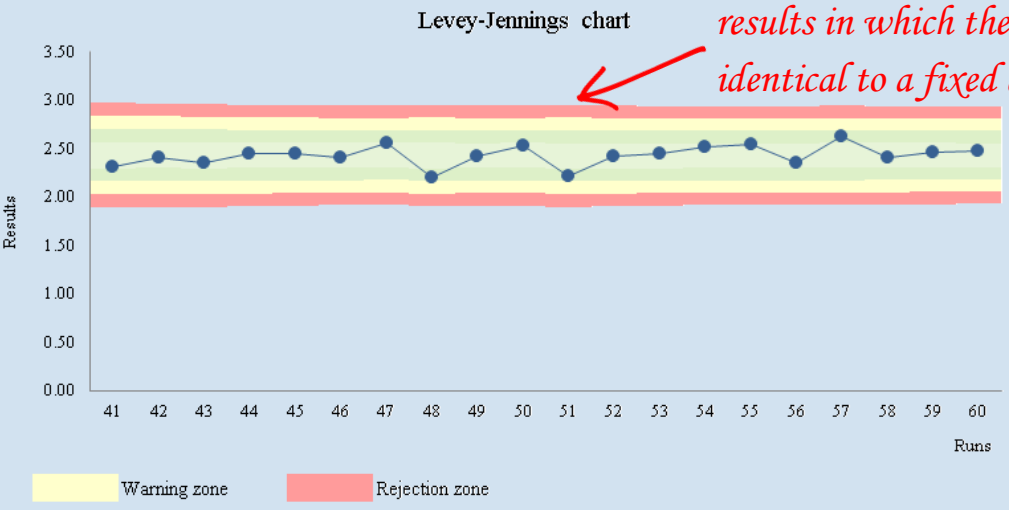


Illustration of a test with stable control results in which the moving average is identical to a fixed average



M. IQC / EQA: Internal Quality Control (Qualitative/Binary Data)

"Pure" qualitative results not quantifiable

Binary results entrance

True and false results ratios

Test:	Immunoassay			Units:	Binary					
Stats:	TP = 95.0%		FN = 5.0%		TN = 95.0%		FP = 5.0%			
no.	Date	Time	Technician	Reag. kit no.	D1	False negative?	D0	False positive?	Material	Lot. no
1	5/12/2015	2:52:24 PM	PPereira	12345ABC67	+	No	-	No	M1	M1234
2	5/13/2015	3:52:24 PM	PPereira	12345ABC68	+	No	-	No	M1	M1234
3	5/14/2015	4:52:24 PM	PPereira	12345ABC69	+	No	-	No	M1	M1234
4	5/15/2015	5:52:24 PM	PPereira	12345ABC70	+	No	-	No	M1	M1234
5	5/16/2015	6:52:24 PM	PPereira	12345ABC71	+	No	-	No	M1	M1234
6	5/17/2015	7:52:24 PM	PPereira	12345ABC72	+	No	-	No	M1	M1234
7	5/18/2015	8:52:24 PM	PPereira	12345ABC73	+	No	-	No	M1	M1234
8	5/19/2015	9:52:24 PM	PPereira	12345ABC74	+	No	-	No	M1	M1234
9	5/20/2015	10:52:24 PM	PPereira	12345ABC75	+	No	+	YES	M1	M1234
10	5/21/2015	11:52:24 PM	PPereira	12345ABC76	+	No	-	No	M1	M1234
11	5/22/2015	12:52:24 AM	PPereira	12345ABC77	+	No	-	No	M1	M1234
12	5/23/2015	1:52:24 AM	PPereira	12345ABC78	+	No	-	No	M1	M1234
13	5/24/2015	2:52:24 AM	PPereira	12345ABC79	+	No	-	No	M1	M1234
14	5/25/2015	3:52:24 AM	PPereira	12345ABC80	-	YES	-	No	M1	M1234
15	5/26/2015	4:52:24 AM	PPereira	12345ABC81	+	No	-	No	M1	M1234
16	5/27/2015	5:52:24 AM	PPereira	12345ABC82	+	No	-	No	M1	M1234
17	5/28/2015	6:52:24 AM	PPereira	12345ABC83	+	No	-	No	M1	M1234
18	5/29/2015	7:52:24 AM	PPereira	12345ABC84	+	No	-	No	M1	M1234
19	5/30/2015	8:52:24 AM	PPereira	12345ABC85	+	No	-	No	M1	M1234
20	5/31/2015	9:52:24 AM	PPereira	12345ABC86	+	No	-	No	M1	M1234

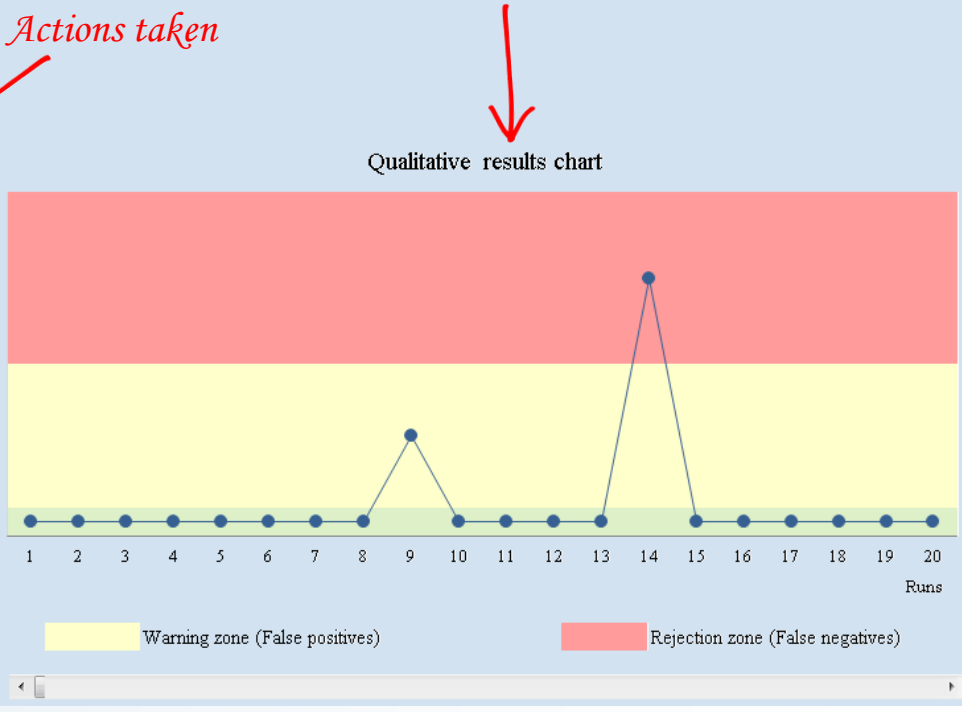
Rejection or warning notice

Rule violation	Excl	Actions
Warning		CAPA 72/2019
Reject		CAPA 87/2019

Possibility of rejection of results

Actions taken

Illustration of a test with accepted results and results in the warning and rejection zones



N. IQC / EQA: External Quality Assessment

PT / EQA scheme for numerical results of ordinal tests

z-value computation

Test:	Immunoassay			Units:	A		Legend (z):	≤2	
Bias ave	7.2%			s_R	8.1%		]2, 2[	≥3	
Exercise	Ref result	Lab result	Bias	s_group		No. of labs	z	Performance	Actions
1	240.000	240.000	0.0%	6.0%	14.400	12	0.000	Satisfactory performance	
2	242.000	239.000	-1.2%	7.0%	16.940	17	-0.177	Satisfactory performance	
3	243.000	242.000	-0.4%	9.0%	21.870	14	0.046	Satisfactory performance	
4	240.000	300.000	25.0%	10.0%	24.000	15	2.500	Questionable performace	
5	239.000	240.000	0.4%	9.0%	21.510	19	0.046	Satisfactory performance	
6	242.000	242.000	0.0%	8.0%	19.360	11	0.000	Satisfactory performance	
7	242.000	325.000	34.3%	8.0%	19.360	10	4.287	Unsatisfactory performance	CAPA 102/2019
8	242.000	241.000	-0.4%	8.0%	19.360	12	-0.052	Satisfactory performance	
9									

Actions taken

PT / EQA scheme for binary results of ordinal or nominal tests

Test:	Immunoassay			Units:	Binary	
Exercise	Ref result	Lab result	Agreement?	No. of labs	Performance	Actions
1	+	+	Yes	12	Satisfactory performance	
2	-	-	Yes	17	Satisfactory performance	
3	+	+	Yes	14	Satisfactory performance	
4	-	+	NO	15	Unsatisfactory performance	CAPA 88/2019
5	+	+	Yes	19	Satisfactory performance	
6	+	-	NO	11	Unsatisfactory performance	CAP 109/2019
7	+	+	Yes	10	Satisfactory performance	
8	-	-	Yes	12	Satisfactory performance	
9						

Conformity assessment

Actions taken

Quality Control of Qualitative Tests for Medical Laboratories

Paulo Pereira, Ph.D.

ISBN 978-989-20-9857-9

Depósito Legal n.º 462249/19

Paulo Pereira, Ph.D., Lisbon, Portugal

e-mail: pauloarpereira@outlook.com

Internet, training courses and for ordering:

http://medlabquality.com

Cover: “Emergir da luz” (emerge from the light) is

a painting by Gracinda de Sousa, 1999

Printed in Portugal by Grafisol

In memory of

Maria Gracinda Gaspar de Sousa, M.D.

(b. December 26, 1952,

d. July 13, 2018)

This page intentionally left blank

Preface

"Everything you always wanted to know about Quality Control (but were afraid to ask)."

This phrase, borrowed from the famous Woody Allen's movie and David Reuben's book, on which it was based, was what I first sensed when the author showed me this book.

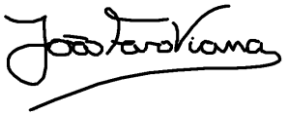
In fact, having only experienced the classical tools of Quality Control applied to quantitative data, and feeling puzzled by the "Uncertainty vs. Total Error" debates, the book filled many of my knowledge gaps on the practical use of Quality Control and Quality Assurance in the Medical Laboratory.

The author's background in a Blood Bank Laboratory explains the choice of the contents, and his teaching experience explains how they are structured, but only his passion for these subjects, that the reader will undoubtedly feel, can explain the thoroughness of his approach.

Besides the contents, the reader will find many practical calculation tools in Excel®. There is, of course, a lot of software available on the market but, for me, exploring these tools will be the best way for the reader to learn and acquire more deeply the otherwise not so easy concepts presented herein.

Lisbon, October 7, 2019

João Faro Viana, M.D.



This page intentionally left blank

Acknowledgments

I have to start by thanking my awesome wife, Sandra, for loving and believing in me, for reading early drafts, giving me advice about the cover, and for all the time spent on editing. She was as important to get this book done as I was. Thank you so much, dear. My thanks also go to my sweet daughters, Madalena and Maria, whose constant lovely and goofy smiles gave me all the more motivation!

I would like to thank João Faro Viana, M.D., who wrote the preface and significantly contributed to the book turning out as it is due to his careful review.

I would also like to thank Teresa Cruz, MSSc, for proofreading the text.

I would like to acknowledge all the people who motivated me and gave me the scientific support to develop my skills in quality control in the med lab: Gracinda de Sousa, M.D., Maria João Alpoim, M.Sc., Jorge Pereira, M.D., and Alexandra Ribeiro, B.Sc., Portuguese Institute of Blood and Transplantation, Jerard Seghatchian, M.D., Ph.D., International Consultancy in Blood Components Quality / Safety Improvement, Audit / Inspection and DDR Strategy, Mário Pádua, M.D., Pulido Valente Hospital, James O. Westgard, Ph.D., School of Medicine and Public Health, University of Wisconsin, Pedro Encarnação, Ph.D., School of Business and Economics, Catholic University of Portugal, Bertil Magnusson, Ph.D., SP Technical Research Institute of Sweden and Elvar Theodorsson, M.D., Ph.D., Department of Clinical and Experimental Medicine (IKE), Linkoping University.

Last but not least, I also place on record my sense of gratitude to one and all who, directly or indirectly, helped me in various ways – you know who you are.

Lisbon, October 10, 2019

Paulo Pereira, Ph.D.

This page intentionally left blank

List of abbreviations and acronyms

δ+	Delta Positive
δ-	Delta Negative
ΔRE_{cont}	Random Error Detectable by Quality Control Procedures
ΔSE_{cont}	Systematic Error Detectable by Quality Control Procedures
Δ_{se}	Difference of Sensitivities
Δ_{sp}	Difference of Specificities
A	Absorbance
A_L	Statistic or Representation of the Number of Control Determinations
ALS	Allowable Loss of Sensitivity
ANOVA	Analysis of Variance
ATE	Allowable Total Error
AUC	Area Under the Curve
b	Bias
BIPM	International Bureau of Weights and Measures
C_0	Negative Condition
C_1	Positive Condition
C_5	Weakest Negative Sample Concentration
C_{50}	Cutoff Concentration
C_{95}	Weakest Positive Sample Concentration
CAPA	Corrective-Action / Preventive-Action
CDC	U.S. Centers for Disease Control and Prevention
CFR	U.S. Code of Federal Regulations
ChLIA	Chemiluminescent Immunoassay
CI	Confidence Interval
CITAC	Cooperation on International Traceability in Analytical Chemistry
CLIA	U.S. Clinical Laboratory Improvement Amendments
CLSI	Clinical and Laboratory Standards Institute
CO	Cutoff Signal
$Conc_1$	Concentration of the Starting Solution
$Conc_2$	Concentration of the Final Solution
COI	Cutoff Index
CRM	Certified Reference Material
CV	Coefficient of Variation
D_1	Disease
D_0	Non-Disease
Day_0	Day of Infection
$Day_{pos/ind}$	Difference of the Number of Days Between the Day of the First Indeterminate or Positive Result
df	Degrees of Freedom
DL	Decision Limit
DNA	Deoxyribonucleic Acid
DPMO	Number of Defects per Million Opportunities
E	Efficiency
EA	European Co-operation for Accreditation

EDTA	Ethylenediaminetetraacetic Acid
EIA	Enzyme Immunoassay
ELISA	Enzyme-linked Immunosorbent assay
EQA	External quality Assessment
EURAMET	European Association of National Metrology Institutes
EUROLAB	European Federation of National Associations of Measurement, Testing and Analytical Laboratories
$f(x)$	Mathematical Function
FDA	U.S. Food and Drug Administration
FDR	False Discovery Rate
FN	False-Negative
FNR	False-Negative Rate
FOR	False-Omission Rate
FP	False-Positive
FPR	False-Positive Rate
GLP	Good Laboratory Practices
GMP	Good Management Practices
GUM	Guide to the Expression of Uncertainty in Measurement
HCV	Hepatitis-C Virus
HIV	Human Immunodeficiency Viruses
HL	High Limit of
HLA	Human Leukocyte Antigen
IEC	International Electrotechnical Commission
IFCC	International Federation of Clinical Chemistry and Laboratory Medicine
IPAC	Instituto Português de Acreditação
IQC	Internal Quality Control
ISO	International Organization for Standardization
IT	Information Technology
IUPAC	International Union of Pure and Applied Chemistry
J	Youden's Index
k	Coverage Factor
k pair	Sensitivity and 1-Specificity Pair per Discriminator
L	Control Limits
LL	Low Limit of
LoD	Limit of Detection
m	Number of Replicate Determinations
n	Number of
NA	Negative Agreement
NAT	Nucleic Acid Test
NCCLS	National Committee for Clinical Laboratory Standards
NIST	National Institute of Standards and Technology
OA	Overall Agreement
p_{ed}	Probability of Error Detection
p_{fr}	Probability of False Rejection
p_n	Percentile of n

PA	Positive Agreement
PCP	Proportion of the Number of Individuals with True Condition in the Population
PCR	Polymerase-Chain-Reaction
PDCA	Plan-Do-Check-Act
PDTS	Preliminary Draft Technical Specification
PT	Proficiency Testing
Pr	Prevalence
QC	Quality Control
QMS	Quality Management System
QCM	Quality Control Material
R&D	Research and Development
RIBA	Recombinant ImmunoBlot Assay
RMS	Root Mean Square
RNA	Ribonucleic Acid
ROC	Receiver Operating Characteristic
s	Standard Deviation
s^2	Variance
s_b	Bias Standard Deviation
s_d	Standard Deviation of the Differences of Paired Results
s_I	Intermediate Standard Deviation
s_r	Repeatability Standard Deviation
s_Rw	Within-Laboratory Reproducibility Standard Deviation
S	Sample Signal
S/CO	Signal-to-Cutoff Ratio
se	Sensitivity
$\mathrm{SE}_\mathrm{crit}$	Critical Systematic Error
sp	Specificity
SPC	Statistical Process Control
SYKE	Finnish Environment Institute
TAE	Total Analytical Error
TAT	Turnaround Time
TC	Technical Committee
TN	True Negative
TNR	True Negative Rate
TP	True Positive
TPR	True Positive Rate
TQM	Total Quality Management
TR	Technical Report
TS	Technical Specification
u_b	Bias Uncertainty
u_c	Combined Uncertainty
$u_{c(\mathrm{ref})}$	Reference Value Standard Uncertainty
U	Expanded Uncertainty
UV	Ultraviolet
WHO	World Health Organization

WP	Window-Period
$\overline{x}$	Average
VIM	Vocabulary of International Metrology
Vol_1	Volume of the Starting Solution
Vol_2	Final Volume of the New Solution

Introduction

One of the most famous quotes in business management states that “if you don’t measure, you don’t know, and if you don’t know, you can’t manage.” So, “if you can’ measure it, you can’t improve it.” “I got my Ph.D. by taking a theory, testing it, and then proving my results. Now, proving my results means that I take those results and I turn them over to other scientists to see if they can replicate them, and if they cannot, then my theory was wrong. See, that’s science. A consensus of evidence that we call truth.” These quotations are entirely applicable to laboratory quality management, including quality control.

As a general quality control practice, the laboratorian mostly deals with statistical approaches primarily intended to be used in clinical chemistry, hematology, urine, and toxicology tests expressing quantitative results. The reason for this could be the more substantial use of quantitative experiments in the diagnosis when compared to qualitative ones. The primary focus of attention on the quality control of qualitative results should be to assure the trueness thereof. From a risk-based viewpoint, untrue results are related to a high-risk of wrong clinical decisions. For example, false-negative results in virology tests in a Blood Bank are associated with a high-risk of the receptor of the blood component being infected post-transfusion. Usually, the beginners in the validation of qualitative tests are introduced to the Bayesian statistics to calculate the probabilities of binary results (positive / negative) occurring in specific samples [1]. Whenever possible, quality control practices are similar to those used in quantitative tests. However, this is only possible when a numerical result is available, as in the case of binary results classified in an ordinal scale according to a specific cutoff. Nevertheless, this practice could be viewed as having several limitations, an example of which is the application of the classic “Westgard rules” [2] to the Levey-Jennings charts [3].

Typically, condition accuracy (condition sensitivity and condition specificity), e.g., diagnostic accuracy, is expressed in a single ratio. However, the limitations of this practice should be understood, and the importance of confidence interval computation recognized. It is critical that laboratorians understand the pros and cons of any statistical models for making reliable and consistent (to specifications) decisions.

The most problematical issue in quality control is probably the determination of measurement uncertainty (5.3 of [4],[5]) required primarily by ISO standards. Several myths remain in the medical laboratory field, contributing to a misinterpretation of its application [6]. The chemistry laboratory faced similar problems, but its computation has been harmonized in the empirical methods for over 18 years [7]. The “Uncertainty Approach” (D.5 of [5]) and “Error Approach” (D.4 of [5]) are different visions to the verification of compliance of results. Measurement uncertainty expresses the statistical dispersion of the values attributed to a measured quantity, and Physical and Chemical scientists unanimously consider it as a more representative concept than the “Error Approach.” However, its

application in the med lab is not successful. Its application to binary results based on an ordinal scale, has been published [8], and it focuses on the uncertainty of results close to the cutoff, i.e., the clinical decision value. Therefore, the determination of the "gray zone" using the standard measurement uncertainty is reinforced.

Whenever the "gray zone" is used, the classification of the results in the ordinal scale is trinary (positive / indeterminate / negative). Consequently, a new definition of the seroconversion window period is suggested considering the period until the first indeterminate result (if it happens) instead of the first positive results [9]. This period is viewed as more realistic for specific infected individuals.

Nonetheless, measurement uncertainty can be computed solely on quantitative results. An alternative method is used to determine the uncertainty of qualitative results. Therefore, a novel definition of "condition uncertainty" is suggested using a 95% confidence interval [10]. Its interpretation is close to the expanded uncertainty, i.e., larger confidence intervals represent a lower statistical chance of trueness.

So, why publish a book called "Quality control of qualitative tests for medical laboratories"? Well, the idea for the text began to take shape about three years ago. This is a book written primarily for the laboratorian and aims to substantiate the selection of the best statistical tools considering the intended use of the qualitative tests' results (fitness for purpose). After reading the standards, typically the lab staff poses several questions related to a consistent implementation of the requirements. The purpose of the book is to answer most of these questions in a three-pronged vision: the statistical, the clinical, and the regulatory vision. The reliability of the last two depends on the consistency of the statistical tools for the intended use of the results. The technical requirements are seen as being integrated into quality management systems based on ISO 15189 [11] or ISO 9001 [12] standards.

The book presents an easy-to-read introduction of the principles and of several examples. The laboratorian should have basic statistical skills and know-how in quality control for a more natural interpretation of the approaches.

Study cases are presented for a more practical view of the theoretical approaches. Since there are several types of qualitative tests, the examples presented here do not include all the methods. Although this could be seen as a limitation, statistical tools can be used in most of the qualitative methods.

All the computations can be done using a conventional computer spreadsheet. The reader can, therefore, easily transpose the functions of the spreadsheet file. All functions are compatible with Excel® (Microsoft®, Redmond, Washington, USA) software. Although the robustness of Excel® is often questioned by statisticians, its use for laboratory data treatment, when verified by data comparison, strongly supports the confidence in its results and the subsequent discussions and conclusions. The spreadsheets are intended only for research purposes and to demonstrate the case studies presented. We strongly encourage the

use of commercial software for laboratory results evaluation, which is available anywhere.

An important pre-evaluation action is the verification of any data point that differs significantly from other observations, referred to as "outlier," to avoid misinterpretation. This check is cross-sectional to all statistical tests. Grubbs test [13] is suggested. Whenever normality of data is required (parametric tests), their distribution can be verified by tests such as the D'Agostino's K^2 [14]. See **SpreadsheetsA-OutliersAndNormalityOfDistribution** for examples of data verification.

Harmonized vocabulary is used to be easily recognized by the med lab staff. Part of the terminology is from the Vocabulary of International Metrology (VIM) [15], and some terms intended solely for the medical laboratory are mostly available in the Clinical Laboratory Standards Institute's "Harmonized terminology database" [16]. For instance, the "qualitative" term used in the book title is not part of the Vocabulary of International Metrology (VIM). However, it is immediately recognized in the medical laboratory. The International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) and the International Union of Pure and Applied Chemistry (IUPAC) published a "Vocabulary on nominal property, examination, and related concepts for clinical laboratory sciences" referred to as "IFCC-IUPAC Recommendations 2017" [17]. The ordinal examination results with a quantitative background scale are expressed by "ordinal tests." For example, binary results (positive / negative, reactive / no reactive) classified on an ordinal scale according to a decision point (cutoff). On the other hand, "nominal tests" are those related to qualitative results with no other related expression. For example, agglutination / no agglutination in a slide for blood typing. We have avoided using terminology that is easily understood in general metrology, but not identified in the med lab.

Outline of the book's structure:

Chapter 1 – ISO compliance introduces mainly ISO 15189 for the accreditation of medical laboratory methods or tests. For a consistent application of this global standard, the laboratorian must understand its specifications. We have discussed the use of most of its technical requirements that involve the selection, verification, validation, measurement uncertainty, internal quality control, and external quality assessment / proficiency testing (EQA / PT) of qualitative results. Moreover, we have crossed ISO 15189 with ISO 9001 requirements for a more natural interpretation of this guideline, which is oriented to a generic implementation of a quality management system. How do we meet the referred ISO claims? See the following chapters for suggested methodologies.

Chapter 2 – Significant causes of uncertainty in qualitative tests discusses the main sources of error that can cause untrue binary results. As the test methodology is essential to recognize the most common analytical causes of failure, we have presented a brief overview of qualitative test design. The impact of the

analytical error on the cutoff trueness is discussed, as well as the effect of the analytical error on the accuracy of the classification of binary results. The importance of the "gray zone" and the associated trinary classification to minimize the impact of analytical error in the results is debated. Moreover, the biased results due to biological factors are presented with a focus on the seroconversion window period. Furthermore, the contribution of other possible sources of bias to the lack of representativeness of patients' samples is also pondered. Finally, the impact of interferences in bias is discussed. This debate is important for a better focus on the use of the quality control tools that allow us to see what is and what is not measurable (limitation of the studies).

Chapter 3 – Measurement uncertainty and total analytical error in qualitative methods introduces both the "Uncertainty Approach" and the "Error Approach" (also recognized as "Traditional Approach" or "True Value Approach"). The challenge is to introduce the laboratorian to the similarities and differences of the visions, wherein empirical models are considered for both. While not ignoring the usefulness of the modular models to the manufacturer, they are not discussed further here since they are not meant to be used in medical laboratory practice. The models presented are based on recognized protocols in med lab requiring data from single-laboratory validation, interlaboratory comparisons or EQA /PT. The importance of the metrological traceability of the results is considered. Compliance assessment is associated with the empirical estimate of the "gray zone." Lastly, the evaluation of analyte concentrations near the cutoff is presented as a complementary tool to estimate an identical zone.

Chapter 4 – Performance of binary classification tests is based on condition accuracy, probably the most well-known methodology for validating qualitative results. In this chapter, we introduce the basis of the statistics concepts applied and discuss the importance of the samples to the robustness of the estimates. We have used 2x2 contingency tables, followed by a discussion about the value of the analysis of the numerical data to distinguish between two or more tests with identical condition sensitivity and specificity. Lastly, the window period is presented using a binary and trinary results logic.

Chapter 5 – Agreement of binary classification tests is intended to lead the reader to validation where samples with a true condition are unavailable. The reliability of this model compared to the condition's accuracy is weak. Therefore, since the consistency of the results is dependent on the comparative test performance, its selection should be applied uniquely if the condition is unavailable.

Chapter 6 – Computation of the cutoff for "in-house" and modified tests, as the title refers, applies solely to tests prepared in the laboratory requiring cutoff determination. Again, the importance of the samples is critical, now to a "realistic" cutoff. However, the "realism" of the cutoff does not depend only on the samples but also on the intended use of the results. Usually, false-positive results are better accepted than false-negative ones. The computation of the cutoff by the receiver operating characteristic curve is discussed. Although we have tried to use the most accessible language, it is probably the most complex statistical model presented in this book. However, its principle is simple: it provides the various

condition sensitivities and specificities for all the possible cutoff points. The laboratorian selects the point that meets the requirements related to the intended use of the results, i.e., according to the clinical application. An area ranking allows the classification of the detection capability of the test for a certain cutoff.

Chapter 7 – Internal quality control and external quality assessment / proficiency testing debate models suitable for qualitative tests. The internal quality control principles are discussed to aid the selection of the best designs based on a qualitative logic. The DPMO-derived and SE_{crit}-derived sigma metrics express the capability to meet the specifications. Models are presented for variables using numerical results (ordinal tests), and an application to monitor "pure" qualitative results (nominal tests). Both methodologies are intended to control the loss of sensitivity in the qualitative tests. EQA /PT is introduced.

References

1. Clinical and Laboratory Standards Institute. EP12-A2 User protocol for evaluation of qualitative test performance, Approved guideline. 2nd ed. Wayne (PA): CLSI, 2008.

2. Westgard JO, Barry PL, Hunt MR, Groth T (1981). A multi-rule Shewhart chart for quality control in clinical chemistry. Clin Chem, 27(3):493-501.

3. Levey S, Jennings ER (1950). The use of control charts in the clinical laboratory. Am J Clin Pathol, 20(11):1059-66.

4. Bureau International des Poids et Mesures. JCGM 200 International vocabulary of metrology - Basic and general concepts and associated terms. 3rd ed. (2008 version with minor corrections). Sèvres: The Organization, 2012. Retrieved from: https://www.bipm.org/en/publications/guides/vim.html. Accessed: October 1, 2019.

5. Bureau International des Poids et Mesures. Evaluation of measurement data - Guide to the expression of uncertainty in measurement. JCGM 100:2008, GUM 1995 with minor corrections. Sèvres: BIPM, 2008. Retrieved from: https://www.bipm.org/en/publications/guides/gum.html. Accessed: October 1, 2019.

6. Pereira P, Seghatchian J. Letter to the Editor: Balance of the unsuccessful systematization of measurement uncertainty in medical laboratories. Transfus Apher Sci 2017, 56(2):103-104.

7. EURACHEM/CITAC. Quantifying uncertainty in analytical measurement. 2nd ed. Europe: The Organizations, 2000.

8. Pereira P, Magnusson B, Theodorsson E, Westgard J, Encarnação P. Measurement uncertainty as a tool for evaluating the “grey-zone” to reduce the false negatives in immunochemical screening of blood donors for infectious diseases. Accred Qual Assur 2016, 21(1):25-32.

9. Pereira P, Westgard J, Encarnação P, Seghatchian J. Analytical model for calculating indeterminate results interval of screening tests, the effect on seroconversion window period: a brief evaluation of the impact of uncertain results on the blood establishment budget. Transfus Apher Sci 2014, 51(2):126-131.

10. Pereira P, Westgard J, Encarnação P, Seghatchian J. Evaluation of the measurement uncertainty in screening immunoassays in blood establishments: Computation of diagnostic accuracy models. Transfus Apher Sci 2015, 52(1):35-41.

11. International Organization for Standardization. ISO 15189 Medical laboratories - Requirements for quality and competence. 3rd ed. Geneva: ISO, 2012.

12. International Organization for Standardization. ISO 9001 Quality management systems - Requirements. 5th ed. Geneva: ISO, 2015.

13. Grubbs FE. Sample criteria for testing outlying observations. Ann Math Stat 1950, 21 (1): 27-58.

14. D'Agostino RB, Belanger A; D'Agostino, Jr RB. A suggestion for using powerful and informative tests of normality. Am Stat 1990, 44 (4): 316-321.

15. Joint Committee for Guides in Metrology (2012). International vocabulary of metrology - Basic and general concepts and associated terms. JCGM 200:2012, JCGM 200:2008 with minor corrections. JCGM. Retrieved from: https://www.bipm.org/en/publications/guides/vim.html Accessed: October 1, 2019.

16. Clinical and Laboratory Standards Institute (2019). Harmonized terminology database. Retrieved from: http://htd.clsi.org/. Accessed: October 1, 2019.

17. Nordin G, Dybkaer R, Forsum U, Fuentes-Arderiu X. Vocabulary on nominal property, examination, and related concepts for clinical laboratory sciences (IFCC-IUPAC Recommendations 2017). Pure Appl Chem 2018, 90(5): 913-935. Retrieved from: https://doi.org/10.1515/pac-2011-0613. Accessed: October 1, 2019.

This page intentionally left blank

Table of contents

This page intentionally left blank

Index of terms