Review

Eurachem/CITAC Guide "Assessment of Performance and Uncertainty in Qualitative Chemical Analysis"—A Medical Laboratory Perspective

Paulo Pereira [iD]

R&D Department, Portuguese Institute of Blood and Transplantation, Avenida Miguel Bombarda, 6, 1000-208 Lisbon, Portugal; paulo.pereira@ipst.min-saude.pt

Abstract: *Background*: The performance assessment of tests that express qualitative results in the medical laboratory is of primary importance in characterization, diagnosis, follow-up, and screening. An important contribution to this type of assessment may be the publication of the Eurachem AQA 2021 guide. The text intends to principally discuss the consistency of the subclauses of this guide with ISO 15189 and CLSI EP12-A2. *Methods*: The study involves a literature review within the scope of qualitative tests. *Results*: Tables are used for crossing AQA. with ISO 15189 and CLSI EP12-A2. *Conclusions*: Consistency with ISO 15189 and CLSI EP12-A2 is demonstrated in the study. Introducing "uncertainty of proportion" reflects the necessity of assessing uncertainties when dealing with qualitative results.

Keywords: clinical; CLSI EP12; ISO 15189; qualitative; sensitivity; specificity

1. Introduction

1.1. The Performance Assessment of Qualitative Results in the Medical Laboratory

The quality control principles in the medical laboratory became more systematized from the 1970s onward. This was due to a vast number of publications that transposed and reviewed approaches to chemistry from a clinical perspective. This development mainly focused on tests that express quantitative quantities, such as 20 IU/L of alanine aminotransferase (ALT). The methodologies initially focused on validation and later on internal quality control (IQC) and external quality assessment (EQA). The availability of guides for evaluating the performance of quantitative tests is vast, with the majority, within the medical laboratory, published by the Clinical Laboratory Standards Institute (CLSI) [1], a global agency. CLSI is an international not-for-profit group that develops laboratory standards for use in the medical laboratory stakeholders using volunteers' expertise. Its standards are recognized by medical laboratories, accreditors, government agencies, and manufacturers of *in vitro* diagnostic medical devices (IVDDs) as reference techniques for improving medical laboratory testing.

Aside from measurement uncertainty [2], we can argue that the medical laboratory's specific performance assessment methods comply with the general principles embodied in the relevant standards, such as ISO 15189 [3]. Medical laboratories can use this global standard to develop quality management systems and assess their own competence by accrediting laboratory tests and methods [4]. As a result of current performance assessment methodologies, laboratories can prove that they operate effectively, even complying with ISO/IEC 17025 [5]. Similar considerations can be made of the various national legal regimes and the manner in which IVDDs are regulated. [6,7].

We emphasize the role of the International Organization for Standardization (ISO), which has developed over 24,222 international standards. A total of 167 national standard-setting bodies are members of ISO, an independent, non-governmental international organization. We can confidently say that no other global organization has contributed as

much to harmonizing practices in industry and services, with ISO 15189 being one of the examples in the medical laboratory.

When it comes to tests that express qualitative results, the number of publications is much smaller. For example, compared to quantitative assessments, CLSI has published a small number of documents for qualitative performance assessment, with the central guide being EP12-A2 [8], which is under review. The evaluation methodology in this type of study focuses on Bayesian probability [9], often crossing with epidemiological principles. Concepts such as uncertainty are not usually associated with these assessments. Performance is generally assessed through clinical sensitivity and clinical specificity, which are the proportions of true results for a given condition (5.3 of [8]), e.g., SARS-CoV-2-infected individuals (condition), and the proportion of negative results in the absence of a given condition (clinical specificity), i.e., in healthy individuals. Both are immediately identified as estimates of diagnostic accuracy. Even when the 95% confidence interval (CI) for these proportions is calculated, the evaluation usually focuses solely on the absolute clinical sensitivity and clinical specificity values. A review of this guide can be found anywhere [10].

The COVID-19 pandemic highlighted the importance of diagnostic accuracy assessment in screening tests. The evaluation, like the one exposed by the Foundation for Innovative New Diagnostics (FIND), a global health nonprofit organization based in Geneva, Switzerland, and a World Health Organization (WHO) Collaborating Center for Laboratory Strengthening and Diagnostic Technology Evaluation, allowed us to assess which test is recommended for screening, in addition to differentiating the performance of both real-time polymerase chain reaction (RT-PCR) tests for ribonucleic acid (RNA) detection and tests for anti-IgG/IgM antibodies of SARS-CoV-2 [11].

1.2. Eurachem and CITAC

Eurachem is a European network of organizations to establish a system for the international traceability of chemical measurements and promote good quality practices. The organization provides a forum for discussing common problems and developing an informed and considered approach to technical and policy issues. Eurachem promotes best practices in analytical measurement by producing authoritative guidance within its expert working groups, publishing guides on the Web, and supporting workshops to communicate good practice. The guidance covers technical issues such as measurement uncertainty evaluation, method validation, and EQA [12].

The Cooperation on International Traceability in Analytical Chemistry, abbreviated to CITAC, is a global organization that aims to discuss how analytical activities could be developed to meet the needs, and it has identified a wide variety of issues to be addressed to ensure that analytical measurements made in different countries or at different times are comparable. These range from developing traceable reference materials and methods to harmonizing analytical quality practices [13].

1.3. Eurachem Qualitative Analysis Working Group

At the Eurachem/CITAC workshop in Lucerne in June 2002, a workshop session was held on qualitative analysis and testing uncertainty. That workshop recommended forming a new Eurachem working group to guide the topic; on the basis of the discussion paper presented at that meeting, the Qualitative Analysis Working Group (QAWG) was created. It aims to prepare guidance on the assessment and expression of uncertainty in qualitative analysis and testing and arrange for appropriate publication and promotion of the above guidelines.

The QAWG was responsible for authoring the first edition of the guide "Assessment of performance and uncertainty in qualitative chemical analysis AQA 2021". The approval date comprised 15 participants from Eurachem member organizations and two participants from CITAC members. The guide is a milestone in Eurachem/CITAC publications, as it is the first on qualitative tests in chemistry.

1.4. Eurachem/CITAC Guide "Assessment of Performance and Uncertainty in Qualitative Chemical Analysis"

Qualitative analysis is frequently used in several analytical fields, such as medical laboratories, where this guide draws upon the experiences of these fields. The manual is intended to assist laboratory staff in choosing and implementing methods to assess the quality of qualitative chemical analysis methods and evaluate uncertainties associated with qualitative chemical analysis. Its purpose is to establish the quantitative reliability of a qualitative analysis result.

The Eurachem/CITAC guide takes into account the following types of criteria:

(a) "Quantitative criteria in which a numerical result is used to categorize a test item as belonging to a pre-established class";
(b) "Qualitative criteria such as the presence or absence of a particular feature, color change on a test, etc." (2 of [14]).

However, it is not centered on binary ordinal quantities, such as those with 0/1 or yes/no values that can be ordered by size, such as positive/negative results classified on an ordinal scale according to a clinical decision point/cutoff. It deals mainly with nominal properties, i.e., with values without size, such as agglutination/no agglutination [15].

The guide, after the introduction, presents the types of qualitative analyses. The performance assessment for qualitative analysis is discussed, involving expressions of confidence in qualitative analysis. Afterward, it suggests how to report the qualitative analytical result. Lastly, conclusions and recommendations emerge, as well as six examples, including the performance assessment in detecting SARS-CoV-2 RNA by nucleic acid amplification.

1.5. Rationale and Objectives

The review of this guide is necessary for the medical laboratory, as it is a guide issued by organizations with peer-reviewed solid publications, as evidenced by the vast number of citations. Although the guide is not exclusive to the medical laboratory, it is covered through keywords with ISO 15189, clinical sensitivity, clinical specificity, and uncertainty of proportions. The guide even features a performance assessment example for an RT-PCR test. In other words, part of the target audience of this document is the medical laboratory staff.

The purpose of this review is to answer the following questions:

(a) How important is the guide to performance assessment of qualitative tests in the medical laboratory?
(b) Can the guide's approach satisfy the technical requirements regarding performance assessment?
(c) Does it fit ISO 15189 requirements?
(d) Does it fit CLSI EP12-A2 recommendations?

2. Materials and Methods

The survey was carried out at PUBMED [16], in December 2021, with the following keywords: analysis, assessment, chemical, Eurachem, performance, qualitative, and uncertainty. The data collection process involved two independent reviewers. No review was found for AQA 2021 publication; therefore, this review did not consider eligibility criteria and study selection. Published reviews would not be expected regarding the publication date of the AQA 2021 (17 November 2021) and the writing date of this text (December 2021). Therefore, the risk of a biased study appears to be negligible.

The sources used were the Eurachem website, specifically its pages "About Eurachem", "Eurachem guides", and "Assessment of performance and uncertainty in qualitative chemical analysis". On the CITAC website, we only consulted the "About CITAC" page. The AQA 2021 is freely accessible. ISO 15189 and CLSI EP12-A2 were already part of our portfolio of documents, which are paid for. The "Guidance on performance evaluation of SARS-CoV-2 *in vitro* diagnostic medical devices" by the Medical Device Coordination Group Document of the European Commission is openly accessible on its page [17]. We

also consulted the "Medical Devices Sector" websites of this Commission [6] and that of the US Food and Drug Administration (FDA) "Medical Devices" [7].

3. Results

In the tables, crossings between clauses or subclauses are noted with a cross. We can understand that the crossed clauses or subclauses comprise similar and replicable methodologies in medical laboratory practice. The clauses or subclauses were identified on the basis of their fit for the purpose of the AQA 2021 in the medical laboratory. Crossing with related clauses or subclauses in whole or in part was allowed.

In the ISO standard, only subclauses related to intra-laboratory performance assessment were considered to have the comparison within the same scope. For example, subclauses 5.6.3 "interlaboratory comparisons" and 5.6.4 "comparability of examination results" are somehow out of focus in this study, so they are not discussed.

Considering the application of the AQA 2021 guide in the medical laboratory, we consider it to crosse with the technical requirements of ISO 15189. Table 1 shows this crossing. The AQA 2021 and CLSI EP12-A2 technical guides were a second cross. Table 2 presents this crossover.

Table 1. AQA 2021 vs. ISO 15189 clauses.

ISO 15189	AQA 2021								
	2 Types of Qualitative Analysis	3.2 Quantification of Qualitative Analysis Performance	3.3 Evaluating False Positive and False Negative Rates	3.4 Limit of Detection and Selectivity	4.2 Likelihood Ratio	4.3 Posterior Probability	4.4 Reliability of Metrics	4.5 Uncertainty of Proportions	5 Reporting the Qualitative Analytical Result
5.5.1.2 Verification of examination procedures	X	X	X	X	X	X	X	X	X
5.5.1.3 Validation of examination procedures	X	X	X	X	X	X	X	X	X
5.5.1.4 Measurement uncertainty of measured quantity values								X [1]	

[1] Measurement uncertainty is not an ISO 15189 requirement for tests that do not express a quantitative value.

Table 2. AQA 2021 vs. CLSI EP12-A2 clauses.

CLSI EP12-A2	AQA 2021								
	2 Types of Qualitative Analysis	3.2 Quantification of Qualitative Analysis Performance	3.3 Evaluating False Positive and False Negative Rates	3.4 Limit of Detection and Selectivity	4.2 Likelihood Ratio	4.3 Posterior Probability	4.4 Reliability of Metrics	4.5 Uncertainty of Proportions	5 Reporting the Qualitative Analytical Result
6 Device familiarization and training	X								
7 Evaluation materials	X	X	X	X				X	
8 Bias and imprecision studies									
9 Comparison methods	X	X		X	X	X	X		
10 Data analysis	X	X		X	X	X	X	X	
6 Device familiarization and training	X								

4. Discussion

Table 1 shows a significant crossover between AQA 2021 and ISO 15189. "Verification", in the ISO standard, comprises the performance assessment of validated examination procedures used without modification before being introduced into routine use. It refers to commercialized tests that have already exhausted validation by the manufacturers of *in vitro* diagnostic medical devices and are approved by a notified body, as is the case in the European Union or the US. On the other hand, the "validation" of examination procedures is derived from the following sources: (a) nonstandard methods; (b) laboratory designed or developed methods; (c) standard methods used outside their intended scope; (d) validated methods subsequently modified. Compared to verification, validation involves more complex models, for example, determining the cutoff in an "in-house" test. The verification fundamentally aims to know if the manufacturer's performance is replicable in the laboratory. The deferent clauses can be operationalized through the AQA.

The crossing with CLSI EP12-A2 clauses is shown in Table 2. We can understand that both the CLSI guide and the AQA 2021 aim to operationalize the technical requirements stipulated mainly in subclauses ISO 15189 5.5.1.2 and 5.5.1.3. The mathematical models for determining clinical/diagnostic accuracy, i.e., clinical sensitivity and clinical specificity, are the same as those of the CLSI EP12-A2. As expected, they are based on Bayesian probability [9]. The same is true for other probabilities computed from a 2×2 contingency table, such as positive and negative predictive values. While clinical accuracy is more relevant to the performance assessment of a given qualitative test, predictive values are more important to the physician. While the former is the proportions of true results in samples with a particular condition and without that condition, as mentioned above, the predictive values are the proportion of individual results with a specific condition and without that condition in positive and negative samples, respectively.

Whenever the diagnosis is unknown, it is possible to calculate, alternatively, the agreement of positive and negative results. Similarly to diagnostic accuracy, both are calculated from the results of the contingency table. Mathematical models are similar to sensitivity and specificity. In this case, the ratios are influenced by non-concordant results.

Compliance assessment is one of the most important and least harmonized topics, not being clear to all medical laboratories when validating clinical sensitivity and specificity results. The European Commission has published some guides on the performance evaluation of *in vitro* diagnostic medical devices, such as SARS-CoV-2 [18]. This document includes specimen type, number of samples, and acceptance criteria for the different performance assessment parameters. As a rule of thumb, the medical laboratory must set clinical sensitivity and specificity targets depending on the intended use of its results. Let us consider the blood bank case versus the clinical pathology laboratory. The sensitivity/specificity tradeoff in blood banks favors sensitivity to minimize the risk of false-negative results, implying a high risk of post-transfusion infection. In contrast, in a pathology laboratory, the sensitivity may be lower than 100% because we can retest a patient without causing harm to third parties. Compliance assessment is poorly discussed in AQA 2021 and EP12-A2. Eurachem

The most significant difference in the diagnostic accuracy approaches is probably due to the view introduced by the AQA, which is based on publications by Pereira et al. [19] (4.4.6 of [20]): uncertainty of proportions. The calculation model is the same as published in the CLSI guide EP12-A2 for 95% CI for clinical sensitivity and clinical specificity. A 95% score confidence interval, attributed to Wilson [21], is calculated in both guides. The two-sided 95% CI for sensitivity or specificity must exceed the lower bound criteria. The criteria are easy to compute; a fixed n is considered for each sample type, or the requirements are recalculated according to the n. For example, for $n = 10$ infected samples, the lower bound criterion is 72%, which happens when sensitivity is 100%. On the other hand, if specificity of 95% is acceptable, the lower bound criterion is 88.8%.

The introduction of the term "uncertainty of proportions" can be understood as a milestone; as far as we know, it is the first global guide to address the uncertainty of binary

positive/negative data. This model is easily replicable for other qualitative outcomes such as blood groups and karyotypes. In fact, the concept is similar to the expanded measurement uncertainty, which is also associated with a 95% CI. Thus, a larger interval expressing the uncertainty indicates a lower likelihood of the sensitivity or specificity value being in the population with the epidemiological characteristics of the samples studied, with a 95% confidence and a beta and alpha risk of 5%. We believe that this introduction will demystify the principle of "impossibility of calculation" in qualitative expressions. This myth is most likely because the measurement uncertainty is solely for quantitative expressions. Note that, for example, subchapter ISO 15189 5.5.1.4. does not apply to qualitative test results.

For a clearer understanding of the uncertainty of proportions, let us present an example test for screening for antibodies against the hepatitis C virus (HCV) by immunoassay. The performance assessment study involves 20 samples from patients diagnosed with HCV infection and 80 healthy subjects. The claimed clinical sensitivity and specificity results are 100% and 90%, respectively. For the target uncertainty in these proportions for a 95% CI, lower bound criteria of 84% for sensitivity and 82% for specificity are claimed. We can interpret that false-negative results are not allowable, admitting up to eight false-positive results. Absolute values are 100% and 99% for sensitivity and specificity, respectively. Therefore, the test is valid according to the first criterion. The lower limits of the 95% CI were 84% for sensitivity and 93% for specificity; hence, the probability of true results is lower than what was claimed. Note the importance of the consistency of claimed results with the number of false results allowed and the number of samples tested.

CLSI EP12-A2 presents a qualitative method-precision experiment for measurand concentrations near the cutoff (C_{50}) (8.3 of [8]), recognized as the "C_5-C_{95} interval", which, as we can understand, is harmonized with IVDD manufacturers. This model is rarely used in the medical laboratory. Anyway, it could be important for "in-house" or modified tests since it lets you know the consistency of "high negatives" and "low positives" in true results with 95% confidence (95% trueness). Low positives should report 95% of positive results. This template is not part of the AQA content.

The area under the receiver operating characteristic (ROC) curve (AUC) [22] fundamentally determines the cutoff point in tests during development, based on the clinical sensitivity/specificity tradeoff. This model allows, for hypothetical cutoff values, to know the clinical sensitivity and specificity for each point. The "best" cutoff is chosen according to the intended use of the reported results. In fact, the cutoff is selected on the basis of the performance assessment of each candidate point. For example, in a blood bank, sensitivity is favored over specificity. None of the guides provides a sufficient discussion of their use in the medical laboratory.

Even though measurement uncertainty can be important in calculating the "gray zone" [23] in ordinal qualitative tests, binary results are classified by comparing a numerical result as a function of a clinical decision point (cutoff). Depending on the order relative to the cutoff, it is classified as positive or negative. For example, if we use the signal-to-cutoff ratio (S/CO), where the cutoff is one, positive is equal to or higher than one, and negative is lower. Pereira et al. [24] demonstrated the calculation of measurement uncertainty in ratios close to this decision point. From this uncertainty, the "guard band"/"gray zone" was calculated, in which the results were classified as indeterminate. The importance of a ternary classification depends on the fitness for the purpose/intended use of the reported results. Again, it will be more significant in a blood bank than in a clinical pathology laboratory. The empirical determination of the "gray zone" is not referred to in any of the guides. Previously, Dimech et al. [25] published a measurement uncertainty study of screening immunoassays based on EQA data.

Furthermore, another important performance assessment tool in this type of test is the detection limit. This limit is measured in molecular biology tests, such as RT-PCR. This value is also recognized as "analytical sensitivity", i.e., the value from which we have 95% true positives, identified as a "95% hit rate", e.g., 10^2 target RNA copies per reaction. Its

determination employs probit regression (5.5 of [26]), also called the probit model. It is used to model dichotomous or binary outcome variables. The inverse standard normal distribution of the probability is modeled as a linear combination of the predictors. It is closely related to the logit function and logit model. This model is not covered in EP12-A2. Despite being referred to in the AQA, it is not presented in detail.

Lastly, let us discuss the importance of the delta value in test performance (4.5 of [20,27]). This tool is important when at least two tests have identical performance assessments, e.g., when clinical sensitivity is equal. The delta value answers the following question: "Which of these tests is most likely to report false or indeterminate results?". This question is important, mainly in validating blood components in blood banks or human organs, cells, and tissues. What is intended to be mitigated goes beyond the risk of false results. It also includes the risk of a negative impact on budget and stock. For example, it is recognized in the case of false results as it implies retesting, elimination of blood components, and suspension of blood donors. Delta values are determined separately for individuals with positive and negative conditions, abbreviated as $\delta+$ and $\delta-$, respectively. The results are interpreted as follows: a higher delta value indicates a lower tendency for a test to produce false or indeterminate results in human samples with the same epidemiological prevalence as the study samples. This approach is not covered in any of the guides.

5. Conclusions

The Eurachem/CITAC technical guide "Assessment of performance and uncertainty in qualitative chemical analysis" is suggested to operationalize the ISO subclauses to verify and validate examination procedures. At this point, AQA 2021, open-source, is an alternative to EP12-A2, being consistent with its methodologies. The AQA examples of clinical sensitivity and specificity are easily replicable in the medical laboratory. It would be important for the compliance assessment to be discussed in depth in future reviews of both guides. The introduction of the uncertainty of proportions allows the medical laboratory to assess the uncertainty in this type of test, which is a surplus concerning ISO 15189. Compared to EP12-A2, no limitations emerged when applying the AQA in the medical laboratory. However, both guides have limitations, especially in validation or even developmental studies. They do not include complimentary performance assessment models, such as the cutoff determination, even though the EP12-A2 allows us to assess their consistency with the "C_5-C_{95}" principle. Another missing point is the computation of the "gray zone", which is important in human product banks for transfusion or transplantation, and the delta value.

Considering the similarities with EP12-A2, it is likely that AQA 2021 could serve as another essential guide for medical laboratory stakeholders, such as med labs, regulatory agencies and IVDD manufacturers.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The author declares being a member of the Eurachem Qualitative Analysis Working Group and Technical Expert on the CLSI Document Development Committee on EP12-A3.

References

1. CLSI. Who We Are. 2021. Available online: https://clsi.org/about/ (accessed on 23 March 2022).
2. Pereira, P.; Westgard, J. Letter to the Editor: Balance of the unsuccessful systematization of measurement uncertainty in medical laboratories. *Transfus. Apher. Sci.* **2017**, *56*, 103–104. [CrossRef] [PubMed]
3. *ISO 15189*; Medical Laboratories-Requirements for Quality and Competence. ISO: Geneva, Switzerland, 2012.
4. Sciacovelli, L.; Aita, A.; Padoan, A.; Antonelli, G.; Plebani, M. ISO 15189 accreditation and competence: A new opportunity for laboratory medicine (Perspective). *J. Lab. Precis. Med.* **2017**, *2*, 79. [CrossRef]
5. *ISO/IEC 17025*; Testing and Calibration Laboratories. ISO: Geneva, Switzerland, 2017.

6. European Commission. Medical Devices-Sector. 2021. Available online: https://ec.europa.eu/health/md_sector/overview_en (accessed on 23 March 2022).
7. FDA. Medical Devices. 2021. Available online: https://www.fda.gov/Medical-Devices (accessed on 23 March 2022).
8. Clinical and Laboratory Standards Institute. *EP12-A2 User Protocol for Evaluation of Qualitative Test Performance, Approved Guideline*, 2nd ed.; CLSI: Wayne, PA, USA, 2008.
9. Bayes, T.; Price, R. *An Essay towards Solving a Problem in the Doctrine of Chance. By the Late Rev. Mr. Bayes, Communicated by Mr. Price, in a Letter to John Canton, AMFRS*; Philosophical Transactions of the Royal Society of London: London, UK, 1763; Volume 53, pp. 370–418.
10. Appold, K. Guideline Promotes Uniformity in Performance Assessment of Qualitative Tests (Lab Guidelines & Standards). *Labmedicine* **2008**, *39*, 273–274.
11. FIND. SARS-CoV-2 Diagnostic Pipeline. 2021. Available online: https://www.finddx.org/covid-19/pipeline/ (accessed on 23 March 2022).
12. Eurachem. A Quick Reference. 2020. Available online: https://www.eurachem.org/index.php/mnu-about (accessed on 23 March 2022).
13. CITAC. About CITAC. 2020. Available online: https://www.citac.cc/general-info/ (accessed on 23 March 2022).
14. Eurachem/CITAC. *CITAC Guide: Assessment of Performance and Uncertainty in Qualitative Chemical Analysis*; Bettencourt da Silva, R., Ellison, S.L.R., Eds.; Eurachem/CITAC, 2021. Available online: https://www.eurachem.org/index.php/publications/guides/performance-and-uncertainty-in-qualitative-analysis (accessed on 23 March 2022).
15. Nordin, G.; Dybkaer, R.; Forsum, U.; Fuentes-Arderiu, X.; Pontet, F. Vocabulary on nominal property, examination, and related concepts for clinical laboratory sciences (IFCC-IUPAC Recommendations 2017). *Pure Appl. Chem.* **2018**, *90*, 913–935. [CrossRef]
16. National Library of Medicine. 2021. Available online: https://pubmed.ncbi.nlm.nih.gov/ (accessed on 23 March 2022).
17. European Commission. Guidance-MDCG Endorsed Documents and Other Guidance. 2021. Available online: https://ec.europa.eu/health/medical-devices-sector/new-regulations/guidance-mdcg-endorsed-documents-and-other-guidance_en (accessed on 23 March 2022).
18. European Commission. Medical Devices-Sector, MDCG 2021–21 Rev.1 Guidance on Performance Evaluation of SARS-CoV-2 *In Vitro* Diagnostic Medical Devices. 2021. Available online: https://ec.europa.eu/health/document/download/729f09dc-9f95-40b9-a62a-a0e9fff1d252_en?filename=mdcg_2021-21_en.pdf (accessed on 23 March 2022).
19. 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*, 35–41. [CrossRef] [PubMed]
20. Pereira, P. *Quality Control of Qualitative Tests for Medical Laboratories*; Paulo Pereira: Lisbon, Portugal, 2019.
21. Wilson, E.B. Probable inference, the law of succession, and statistical inference. *JASA* **1927**, *22*, 209–212. [CrossRef]
22. Swets, J.A. ROC analysis applied to the evaluation of medical imaging techniques. *Invest. Radiol.* **1979**, *14*, 109–121. [CrossRef] [PubMed]
23. Kisner, H.J. The gray zone. *Clin. Lab. Manag. Rev.* **1998**, *12*, 277–280.
24. Pereira, P.; Magnusson, B.; Theodorsson, E.; Westgard, J.; Encarnação, P. Measurement uncertainty as a tool for evaluating the 'gray zone' to reduce the false negatives in immunochemical screening of blood donors for infectious diseases. *Accredit. Qual. Assur.* **2016**, *21*, 25–32. [CrossRef]
25. Dimech, W.; Francis, B.; Kox, J.; Roberts, G. Calculating uncertainty of measurement for serology assays by use of precision and bias. *Clin. Chem.* **2006**, *52*, 526–529. [CrossRef] [PubMed]
26. Clinical and Laboratory Standards Institute. *EP17-A2 Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures, Approved Guideline*, 2nd ed.; CLSI: Wayne, PA, USA, 2008.
27. Crofts, N.; Maskill, W.; Gust, I. Evaluation of enzyme-linked immunosorbent assays: A method of data analysis. *J. Virol. Methods* **1988**, *22*, 51–59. [CrossRef]