OPEN

Belgian Recommendations for Analytical Verification and Validation of Immunohistochemical Tests in Laboratories of Anatomic Pathology

Hannelien Verbeke, PhD, Donald Van Hecke, BSc,† Caroline Bauraing, MSc,‡ Anne Marie Dierick, MD,* Orphal Colleye, PhD,§ Ignace Dalle, MD,† Kathleen Dewachter, MSc,‖ Yves Guiot, PhD,¶ Raphael Lequeu, MSc,‡ Nancy Vanderheyden, MSc,# Karen Zwaenepoel, PhD,** and Romaric Croes, MD††*

Abstract: Analytical verification and validation of immunohistochemical (IHC) tests and their equipment are common practices for today's anatomic pathology laboratories. Few references or guidelines are available on how this should be performed. The study of Sciensano (the Belgian national competent authority regarding licensing of medical laboratories) performed in 2016, demonstrated a significant interlaboratory variation in validation procedures of IHC tests among Belgian laboratories. These results suggest the unavailability of practical information on the approach to the verification and validation of these tests. The existing Belgian Practice Guideline for the implementation of a quality management system in anatomic pathology laboratories has been reviewed to meet this demand and, in addition, to prepare the laboratories for the EU—IVD revised regulations (IVDR). This paper describes Belgian recommendations for the verification and validation of IHC tests before implementation, for ongoing validation, and for revalidation. For each type of test (according to the IVDR classification and the origin) and its intended use (purpose), it addresses how to perform analytical verification/validation by recommending: (1) the number of cases in the validation set, (2) the performance characteristics to be evaluated, (3) the objective acceptance criteria, (4) the evaluation method for the obtained results, and (5) how and when to revalidate. A literature study and a risk analysis taking into account the majority of variables regarding verification/validation of methods have been performed, resulting in an expert consensus recommendation that is a compromise among achievability, affordability, and patient safety. This new consensus recommendation has been incorporated in the aforementioned ISO 15189:2012–based Practice Guideline.

Key Words: immunohistochemical test, CE-IVD test, laboratory-developed test, verification, validation

(*Appl Immunohistochem Mol Morphol* 2024;32:1–16)

The Belgian Commission of anatomic pathology (AP) provided a Practice Guideline, based on the International Standard ISO 15189:2012, to help Belgian laboratories of AP improve their quality management system (QMS).[1] To monitor the implementation of a QMS and the application of the Practice Guideline, Sciensano, the Belgian national competent authority regarding licensing of medical laboratories, performs documentary audits and on-site inspections. As clinicians and patients rely on accurate diagnostic testing, a documentary study has been performed on the validation procedures for test methods applied in Belgian AP laboratories.[2] The insufficient compliance with existing guidelines and recommendations, and even with the ISO 15189:2012 standard, implied the need for univocal and standardized instructions regarding the validation of immunohistochemical (IHC) tests.[2–15] These have recently been incorporated in a revised version of the Practice Guideline.[1]

Despite the existence of guidelines on the validation qof IHC biomarker assays, specific details regarding its application in individual laboratories are lacking.[3,7,9–13,16–19] In daily practice, AP laboratories perform IHC stainings using Conformité Européenne (CE)-marked In Vitro Diagnostic (IVD) (CE-IVD) tests according to the Instructions For Use (IFU) or modified (not according to the IFU), tests for Research Use Only (RUO), and in house developed tests. The

Received for publication February 16, 2023; accepted September 10, 2023.

From the *Department of Quality of laboratories, Sciensano; ¶Laboratory of Anatomic Pathology, Université Catholique Louvain, Brussels; †Laboratory of Anatomic Pathology, AZ St. Lucas, Bruges; ‡Institute of Anatomic Pathology and Genetics, Gosselies; §Laboratory of Anatomic Pathology, CHU Liège, Liège; ‖Laboratory of Anatomic Pathology, AZ St. Maarten, Mechelen; #Laboratory of Anatomic Pathology, UZ Leuven, Leuven; **Laboratory of Anatomic Pathology, UZ Antwerp, Antwerp; and ††Laboratory of Anatomic Pathology, AZ St. Blasius, Dendermonde, Belgium.

H.V. and D.V.H. contributed equally.

The authors declare no conflict of interest.

Reprints: Hannelien Verbeke, PhD, Belgian Scientific Institute of Public Health, Sciensano, Brussels 1050, Belgium (e-mail: hannelien.verbeke@sciensano.be).

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.appliedimmunohist.com.

CE-IVD tests used according to the IFU have to comply with the new European (EU) In Vitro Diagnostic Medical Devices regulation 2017/746 (IVDR) effective since May 26, 2022.[20] All other types of tests can only be used under certain conditions as described in article 5(5) of the IVDR.[20] The absence of a definition for modifications of CE-IVD tests resulting in laboratory-developed tests (LDTs) and the lack of information on how to demonstrate its clinical and analytical performance are 2 major concerns for the laboratories. To clarify this, instructions have been elaborated for initial analytical verification/validation, revalidation, and ongoing verification/validation for each type of test. It also takes into account the requirements of the new IVDR. A literature study and a risk analysis have been performed to obtain this expert group consensus recommendation that is a compromise among achievability, affordability, and patient safety.

DEFINITIONS

IHC test—an IHC staining using a specific primary antibody (Ab), detection method, and staining platform.

IHC biomarker assay—the overall of different IHC tests/stainings.

Verification—confirmation through the provision of objective evidence that specified requirements are fulfilled. These specified requirements have been initially validated by the manufacturer for a specific intended purpose.[4,12,15,21] The laboratory verifies for the same intended purpose and using the same qualified equipment, reagents, and staining protocol (in its own environment, with its own conditions, and employees), the predefined performance characterizes (ie accuracy and repeatability/reproducibility) of the test to demonstrate that they are capable for achieving the same required performance before actual use.[4,12,15,21] This type of verification is also called technical or analytical verification.[12]

Validation—demonstration by examination and provision of objective evidence that the performance characteristics meet the predefined criteria or specified requirements for a specific intended use or application.[4,12,21]

Purpose—the intended use of the test at the time the test was developed.[10]

Accuracy—degree of concordance with a gold standard [eg, reference staining, external quality assessment (EQA) result, non-IHC test].[13]

Analytical sensitivity—the smallest amount of substance in a tissue sample that can accurately be detected by an IHC biomarker assay or limit of detection study.[13]

Diagnostic sensitivity—the probability that a test outcome is positive in a diseased person or the proportion of true positive in patient cases.[13,22]

Analytical specificity—the ability of an IHC biomarker assay to detect one particular substance in a tissue sample, rather than others.[13]

Diagnostic specificity—the probability that a test outcome is negative in a nondiseased person or the proportion of true-negative in healthy cases.[13,22]

Repeatability—the ability of a test to produce the same result for a given tissue sample when repeatedly tested in the same run (intrarun).

Reproducibility—the ability of a test to produce the same result for a given tissue sample when repeatedly tested over different runs (interrun).[13]

Robustness—the ability of a test to remain unaffected by small but deliberate variations in procedural parameters or test conditions.[6,13] Typical variables in a laboratory of AP include a delay to fixation, fixation time, fixative type, pretreatment processing (eg, antigen retrieval), section thickness, antigen stability, reagent stability, and environmental conditions (eg, different operators and equipment).

Readout precision—the ability to reproduce the intraobserver and interobserver readout on a given tissue sample.[13]

TYPES OF TESTS

Execution of an analytical verification or validation is determined by (1) the purpose of the test related to the intended use, (2) the IVDR classification, and (3) the origin of the test.

(1) IHC tests for diagnostic purposes are classified as type 1, and prognostic and pharmaco-predictive IHC tests are, respectively, classified as type 2a and type 2b.[10,23]
(2) At the European level (cf. IVDR) in vitro diagnostic tests are divided into 4 classes, A, B, C, and D, taking into account their intended purpose and inherent risks.[20] IHC tests are classified as class C (except for the detection of infectious agents, classified as class B; *cf.* IVDR article 47 and annex VIII).[20]
(3) The origin of a test can be CE-IVD, modified CE-IVD, for RUO, and in-house developed. The IVDR makes a distinction between IVD tests "CE-marked" and "not CE-marked". LDTs are not clearly defined within the IVDR, but are described as "devices manufactured and used only within health institutions established in the Union."[20]

In compliance with the IVDR, CE-IVD kits should be the first choice.[20] Only when no suitable commercially CE-IVD kit is available for the intended use, a deviation will be accorded on the condition that all requirements as stated in the IVDR are met.[20] This deviation allows the use of LDTs if a laboratory meets the requirements of the standard ISO 15189 or appropriate national requirements, such as the Belgian Practice Guideline. Nonetheless, even if CE-IVD kits are used according to the IFU and for the same purpose, initial testing results may show that the manufacturer's specifications are not being met, forcing the laboratory to modify one or more prescribed specifications of the IFU for optimization of the test.

To clarify the definition and use of LDTs, we propose to subcategorize LDTs as follows:

- Modified CE-IVD test with reference—the modifications do not fall within the range of the IFU, for example, a change in retrieval solution, primary Ab incubation time, and primary Ab dilution factor. However, a reference for

the modified method used (ie, staining protocol) and its specific intended use is available (eg, EQA results, scientific publication, verification/validation data of the test from another laboratory, an optimized staining protocol proposed by an EQA provider, or the manufacturer's application specialist). In this case, a limited validation (in accordance with verification) would be sufficient to confirm the performance specifications as indicated by the reference (Table 1).

- Modified CE-IVD test without reference—the modifications made to the staining protocol deviate from the IFU or the specified intended use and no references are available. In this case, a validation is recommended (Table 1).
- Non–CE-IVD test with reference—RUO products are not intended for diagnostic purposes but are nevertheless used because, for example, the lack or inferior quality of an appropriate CE-IVD Ab /kit. For RUO tests with reference, a simplified validation is proposed as discussed further.
- Non–CE-IVD test without reference or in-house developed test—For non–CE-IVD tests, the clinical performance has to be demonstrated besides the general safety and performance requirements (*cf.* IVDR) and an extensive validation has to be performed.[20] "Clinical performance" means the ability of a device to yield results that are correlated with a particular clinical condition or a (patho)physiological process.

VERIFICATION OR VALIDATION REQUIREMENTS

Verification or Validation Plan and Report

The necessity to verify or validate is often related to the introduction of a new staining platform, a new method (ie, detection system), and/or a new test (ie, Ab). Establishing a plan before execution is recommended. This plan describes the steps to prepare and execute the verification/validation process. In addition, a verification/validation report has to contain the examination results and all data that may influence these results (Supplemental Table, Supplemental Digital Content 1, Supplemental Digital Content 1, http://links.lww.com/AIMM/A427, data recorded in a verification/validation plan and report). Establishing a template for a verification/validation plan and report is useful to record the data in a harmonized manner.

Matrix and Sample Selection

The most appropriate validation set contains the same matrix as in daily practice.[3,4] The matrix refers to: (1) the fixative and processing method (eg, formalin-fixed, paraffin-embedded tissue) and (2) the type and origin of tissue and cytologic specimens.[3] CE-IVD tests have been validated by the manufacturer on a specific matrix for a particular intended use. When the CE-IVD test is used on different matrices (eg, HER2 testing algorithm for breast cancer vs gastric cancer, cell blocks derived from cytologic material) an additional verification or validation of each matrix is recommended. In case of analytical tests applied on matrices without validation (eg, limitations in sample availability), a disclaimer should be included in the patient report that the results should be interpreted with caution.[3,24]

The samples used during optimization and verification are generally quality control materials and selected patient specimens. Control material consists of (archived or residual) tissues or cytologic specimens previously tested by a validated laboratory method (either within or by another laboratory), with known specifications and a similar matrix as the patient specimens of the daily practice.[3,5] Besides control material and patient specimens, reference standards, cell lines, and excess tissue from previous EQA or internal laboratory comparison (peer review) programs may be selected.[4,5] However, matrix-induced effects, including specimen origin, fixative type, fixation time, tissue processing, and antigen retrieval method, have to be taken into consideration as they may influence the test results and conclusions.[4,5,25] Therefore, we recommend selecting as many patient specimens as possible in addition to samples from other sources, even though the latter is useful to determine the robustness of the test.[3,9]

The validation set should include high and low protein expression levels (close to the detection threshold) and should cover the expected range of clinical results.[3,4] Tissue microarrays and multitissue blocks can be used.[3,9,11] They contain multiple previously tested positive

TABLE 1. The Performance Characteristics to be Verified or Validated in Function of the Origin of the Test

	Verification	Limited validation	Validation		
Performance characteristics*	**CE-IVD test**	**Modified CE-IVD with reference**	**Modified CE-IVD without reference**	**Non–CE-IVD (RUO) with reference**	**Non–CE-IVD without reference (in-house developed)**
Accuracy	x	x	x	x	x
Repeatability and reproducibility	x	x	x	x	x
Analytical sensitivity	—	—	x	x	x
Analytical specificity	—	—	x	x	x
Robustness	x†	x†	x†	x†	x†
Readout precision	(x)	(x)	(x)	(x)	(x)

*For the definitions of the performance characteristics, see the section "Definitions" in the main text.
†The performance of a risk analysis can be a useful tool to determine which parameters can be verified in the context of the robustness.
(x): mainly applicable to semiquantitative testing.
RUO indicates Research Use Only.

TABLE 2. Composition of the Validation Set and Determination of the Acceptance Criteria for Verification and Validation of IHC Type 1 Tests

	CE-IVD or modified with reference			
Performance characteristic*	**With experience†**	**Without experience**	**Modified CE-IVD without reference or non–CE-IVD with reference**	**Non–CE-IVD without reference‡**
Accuracy§				
Panel Ab 1 to 15				
T×S‖	$5t^+$ $5t^-$	$10t^+$ $10t^-$	$15t^+$ $15t^-$	$40t^+$ $40t^-$
AC¶ (%)	≥90	≥90	≥93	≥95
From 16th Ab				
T×S	$2t^+$ $2t^-$	$2t^+$ 2t−	3 to $5t^+$ 3 to $5t^-$	NA
AC (%)	≥75	≥75	≥90	NA
Analytical sensitivity				
Panel Ab 1 to 15				
T×S	NA	NA	$15t^+$ $15t^-$	$40t^+$ $40t^-$
AC (%)	NA	NA	≥93	≥95
From 16th Ab				
T×S	NA	NA	3 to $5t^+$ 3 to $5t^-$	NA
AC (%)	NA	NA	≥90	NA
Analytical specificity				
Panel Ab 1 to 15				
T×S	NA	NA	$15t^+$ $15t^-$	$40t^+$ $40t^-$
AC (%)	NA	NA	≥93	≥95
From 16th Ab				
T×S	NA	NA	3 to $5t^+$ 3 to $5t^-$	NA
AC (%)	NA	NA	≥90	NA
Repeatability				
Panel Ab 1 to 15				
T×S	3×3 (1Ab)	3×3 (1Ab)	3×3 (3 Ab)	9×3 (each Ab)
AC (%)	≥90	≥90	≥90	≥90
From 16th Ab				
T×S	NA	NA	NA	NA
AC (%)	NA	NA	NA	NA
Reproducibility				
Panel Ab 1 to 15				
T×S	3×3 (1Ab)	3×3 (1Ab)	3×3 (3 Ab)	9×3 (each Ab)
AC (%)	≥90	≥90	≥90	≥90
From 16th Ab				
T×S	NA	NA	NA	NA
AC (%)	NA	NA	NA	NA
Robustness				
Panel Ab 1 to 15				
T×S	NA	NA	NA	NA
AC (%)	NA	NA	NA	NA
From 16th Ab				
T×S	NA	NA	NA	NA
AC (%)	NA	NA	NA	NA
Others: I/U/S/B/C				
Panel Ab 1 to 15				
T×S	$5t^+$ $5t^-$	$10t^+$ $10t^-$	$15t^+$ $15t^-$	$40t^+$ $40t^-$
AC (%)	≥X#	≥X	≥X	≥X
From 16th Ab				
T×S	$2t^+$ $2t^-$	$2t^+$ $2t^-$	3 to $5t^+$ 3 to $5t^-$	NA
AC (%)	≥X	≥X	≥X	NA

*For the definitions of the performance characteristics, see the section "Definitions" in the main text.

†Laboratories should retrospectively and objectively demonstrate their experience for at least 8 of the 15 antibody biomarkers within the verification panel (2 antibodies with cytoplasmic staining pattern, 2 antibodies with nuclear staining pattern, 2 antibodies with membranous staining pattern, and 2 antibodies with a more complex staining protocol). Criteria to demonstrate experience with the type 1 immunohistochemical method are, for example, at least good or optimal external quality assessment results for

the most recent external quality assessment program within the past 7 years without any change in staining pattern (antibody-antigen interaction) and interpretation of the results (eg, scoring algorithm), traceable internal quality control results over a certain period (at least 6 mo) or no registration of nonconformities with regard to the method/technique used within the past year.

‡For non–CE-IVD type 1 tests without reference and for which limited (control) materials are available, but used within the context of a verified/validated CE-IVD detection method, we recommend starting the initial validation with a set of 2 positive and 2 negative tissues and progressively increase the validation set over time until the target of 40 positive and 40 negative tissues is reached.

§When a type 1 CE-IVD immunohistochemical test is used on a matrix for which the test has not been validated by the manufacturer (eg, cytologic specimens and fresh frozen/nonfixated specimens), we recommend demonstrating the accuracy of the test by staining 3 to 5 positive and 3 to 5 negative tissues depending on the sample availability, provided that the accuracy of the immunohistochemical method has been demonstrated using a selected group of antibodies for the immunohistochemical method verification. When the accuracy has not yet been demonstrated in the context of the immunohistochemical method verification, 10 positive and 10 negative tissues have to be selected.

‖T = number of tissues/samples and S = number of sections/slides of each selected tissue/sample.

¶Acceptance criterion, concordance percentages between expected and obtained results.

#Predefined minimum total score.

Ab indicates antibody; AC, acceptance criterion; B, background, C, counterstaining; I, intensity; IHC, immunohistochemical; NA, not applicable; S, specificity; U, uniformity.

and negative tissues, allowing the comparison between multiple results within one run, resulting in a time and cost-saving approach.[3,9,11] However, the limited and heterogeneous expression of antigens in certain tissues has to be taken into account and could be a contraindication for use.[3,9]

ANALYTICAL VERIFICATION OR VALIDATION PROCEDURE

Optimization (Technical Calibration)

Before verifying or validating an IHC test, the test needs to be technically calibrated or optimized for optimal performance.[3,23] The purpose is to check whether the test results (within the laboratory environment, processing method, equipment, and employees) meet the predefined criteria as defined by the reference. If the default staining protocol does not fulfill the predetermined criteria, the method will need further fine-tuning. During this process, we suggest adjusting one parameter of the staining protocol at a time to evaluate the effect. If the modifications do not fall within the IFU of the manufacturer, the test is considered as an LDT requiring validation.

Documentation and evaluation of all changes made and the subsequently obtained results (eg, staining quality), allow the selection of the best method to be used and justifies the use of an LDT as well.

For optimization, we recommend to evaluate 2 known positive and 2 known negative control samples. As a tissue may contain both antigen-positive cells and negative internal control cells, the evaluation of 2 control samples (including both positive and negative cells) will be sufficient to optimize most diagnostic IHC tests.

Analytical Verification/Validation

The extent of the analytical verification/validation is related to the type of the test. For each type of test and its intended use, verification/validation instructions, including requirements for the methodology, the sample size, and the performance characteristics, are summarized in Tables 2 and 3 and Figures 1, 2, and 3.

Verification of CE-IVD Immunohistochemical Tests According to Instructions for Use

When a CE-IVD test is applied according to the IFU, a verification of the accuracy and repeatability/reproducibility of the test is sufficient before its implementation in daily practice. Care should be taken to interpret the IFU. Despite clear specifications, some package insert states that optimal conditions may vary and can be determined by the individual laboratory depending on the tissue and staining method (eg, pretreatment, primary Ab dilution, incubation time, addition of an enhancement or blocking step, and detection kit from another manufacturer but same detection method as specified in the Ab package insert).

Other examples of changes within the staining protocol for which the test is still considered as a CE-IVD test are: changing the time of counterstaining (noncritical item) and performing the IHC staining on a device of another manufacturer if there are no requirements regarding the device in the package insert of the Ab.

Initial Verification of CE-IVD Immunohistochemical Type 1 Tests

As an extended verification of all individual, type 1 tests is impractical due to the limited (control) material, time, and expense of the verification process, we recommend performing a verification of the "method" by evaluating the performance characteristics of a group of IHC Ab biomarkers using the same detection method and staining platform, followed by a limited verification of all other individual type 1 tests within this method.[26] For this "method verification," we recommend selecting a panel of 15 Abs representing the diagnostic use, the pretreatment method (eg, different pretreatment buffers), and the staining pattern (ie, membranous, cytoplasmic, and nuclear) as illustrated in Table 1 and Figure 1. For IHC tests with both diagnostic and prognostic/predictive applications, the recommendations for type 2 IHC tests should be followed (section 2.1.2).[3]

To demonstrate the repeatability and the reproducibility of IHC tests, it is recommended to stain at least 3 tissues in triplicate.[5,7,9] In daily practice equipment and test validation overlap implicating that the repeatability/reproducibility of a selection of IHC tests has already been demonstrated during the validation of the staining platform. For that reason, we recommend selecting one Ab

TABLE 3. Composition of the Validation Set and Determination of the Acceptance Criteria for Verification and Validation of IHC Type 2 Tests

	CE-IVD or modified with reference			
Performance characteristic*	**With experience†**	**Without experience**	**Modified CE-IVD without reference or non–CE-IVD with reference**	**Non–CE-IVD without reference**
Accuracy				
Initial matrix				
Type 2a				
T×S‡	$7t^+$ $7t^-$	$15t^+$ $15t^-$	$30t^+$ $30t^-$	$40t^+$ $40t^-$
AC§ (%)	≥90	≥93	≥95	≥95
Type 2b				
T×S	$10t^+$ $10t^-$	$20t^+$ $20t^-$	$30t^+$ $30t^-$	$40t^+$ $40t^-$
AC (%)	≥90	≥95‖	≥95	≥95
Other matrix/origin¶				
T×S	$5t^+$ $5t^-$	$10t^+$ $10t^-$	NA	NA
AC (%)	≥90	≥90	NA	NA
≠FFPE#				
T×S	$3t^+$ $3t^-$	$5t^+$ $5t^-$	NA	NA
AC (%)	≥83	≥90	NA	NA
Analytical sensitivity				
T×S	NA	NA	$30t^+$ $30t^-$	$40t^+$ $40t^-$
AC (%)	NA	NA	≥95	≥95
Analytical specificity				
T×S	NA	NA	$30t^+$ $30t^-$	$40t^+$ $40t^-$
AC (%)	NA	NA	≥95	≥95
Repeatability**				
Initial matrix				
T×S	3×3	3×3	3×3	9×3
AC (%)	≥90	≥90	≥90	≥90
Other matrix/origin				
T×S	3×3††	3×3††	NA	NA
AC (%)	≥90	≥90	NA	NA
≠FFPE				
T×S	3×3††	3×3††	NA	NA
AC (%)	≥90	≥90	NA	NA
Reproducibility**				
Initial matrix				
T×S	3×3	3×3	3×3	9×3
AC (%)	≥90	≥90	≥90	≥90
Other matrix/origin				
T×S	3×3††	3×3††	NA	NA
AC (%)	≥90	≥90	NA	NA
≠FFPE				
T×S	3×3††	3×3††	NA	NA
AC (%)	≥90	≥90	NA	NA
Robustness‡‡				
Initial matrix				
T×S	$5t^+$ $5t^-$	$5t^+$ $5t^-$	$5t^+$ $5t^-$	$10t^+$ $10t^-$
AC (%)	≥90	≥90	≥90	≥95
Other matrix/origin				
T×S	$5t^+$ $5t^-$	$5t^+$ $5t^-$	NA	NA
AC (%)	≥90	≥90	NA	NA
≠FFPE				
T×S	$3t^+$ $3t^-$	$5t^+$ $5t^-$	NA	NA
AC (%)	≥83	≥90	NA	NA

TABLE 3. (*continued*)

Performance characteristic*	CE-IVD or modified with reference: With experience†	CE-IVD or modified with reference: Without experience	Modified CE-IVD without reference or non–CE-IVD with reference	Non–CE-IVD without reference
Others: I/U/S/B/C and/or readout precision§§				
Initial matrix				
T×S	$5t^+$ $5t^-$ (readout precision)	$5t^+$ $5t^-$ (readout precision)	$5t^+$ $5t^-$ (readout precision)	$5t^+$ $5t^-$ (readout precision)
AC (%)	≥90 I/U/S/B/C: >X‖‖	≥90 I/U/S/B/C: >X	≥90 I/U/S/B/C: >X	≥90 I/U/S/B/C: >X
Other matrix/origin				
T×S	$5t^+$ $5t^-$ (readout precision)	$5t^+$ $5t^-$ (readout precision)	NA	NA
AC (%)	≥90 I/U/S/B/C: >X	≥90 I/U/S/B/C: >X	NA	NA
≠FFPE				
T×S	$3t^+$ $3t^-$ (readout precision)	$5t^+$ $5t^-$ (readout precision)	NA	NA
AC (%)	≥83 I/U/S/B/C: >X	≥90 I/U/S/B/C: >X	NA	NA

*For the definitions of the performance characteristics, see the section "Definitions" in the main text.

†Laboratories should demonstrate experience with the test through at least good or optimal external quality assessment results for at least 2 consecutive external quality assessment programs and through concordant inter/intraobserver results.

‡T = number of tissues/samples and S = number of sections/slides of each selected tissue/sample.

§Acceptance criterion, concordance percentages between expected and obtained results.

‖Applicable for HER2, Estrogen Receptor (ER), and Progesterone Receptor (PR). For other type 2 immunohistochemical biomarkers, for example, PD-L1, other acceptance criteria may be applied as reported in national or international publications.

¶Other tissue as initially verified and for which another scoring algorithm is applied for the interpretation of the results (eg, HER2 testing algorithm for breast cancer vs gastric cancer).

#Example: decalcification, cytologic specimen.

**To demonstrate repeatability/reproducibility, it is recommended to test at least one positive, one weakly positive, and one negative specimen in triplicate.

††Verification or validation is not necessary for matrix changes if the staining protocol has not been modified.

‡‡As many variables, such as delay to fixation, fixation time, fixative type, tissue processing, section thickness, antigen stability, reagent stability, and environmental conditions, may affect the robustness of type 2 tests, a verification of this performance characteristic is recommended, based on the results of a risk assessment.

§§When different matrices require different scoring algorithms, a readout precision is performed on each matrix separately.

‖‖predefined minimum total score.

Ab indicates antibody; AC, acceptance criterion; B, background, C, counterstaining; FFPE, formalin-fixed, paraffin-embedded; I, intensity; IHC, immunohistochemical; NA, not applicable; S, specificity; U, uniformity.

biomarker from the mentioned panel to demonstrate the repeatability/reproducibility of the IHC method.

In case of limited availability of appropriate verification/validation samples (eg, rare diseases/tumors), we recommend a stepwise demonstration of the accuracy of the test, allowing temporary approval and implementation in daily practice. After the first staining of 2 positive and 1 negative tissue, the number of specimens is progressively increased to reach 10 positive and 10 negative specimens.

Initial Verification of CE-IVD Immunohistochemical Type 2 Tests

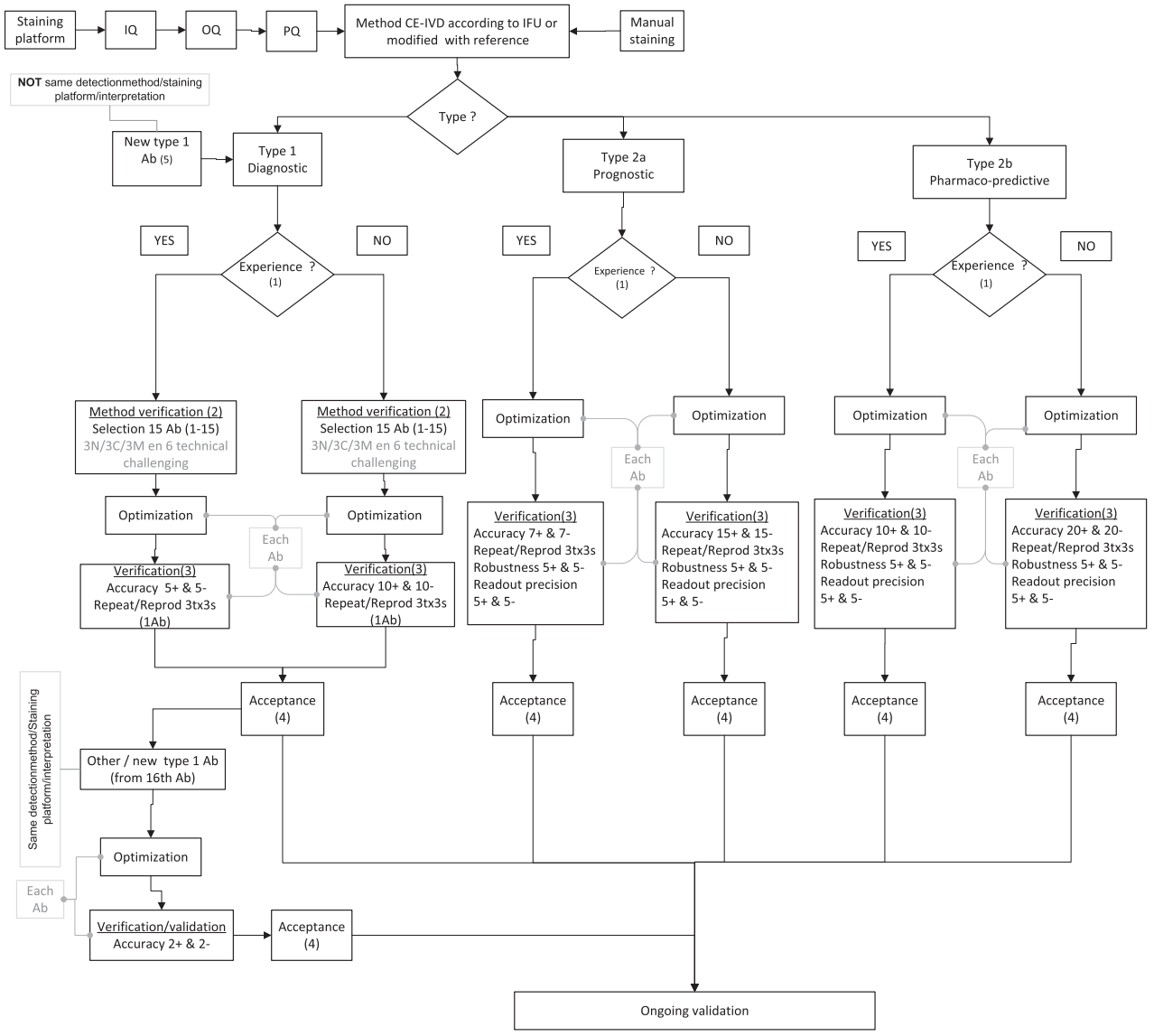
To verify IHC tests with prognostic or predictive purposes, we recommend applying the (inter)nationally published guidelines taking into account the fast-evolving scientific developments (eg, evidence-based medicine). Despite the College of American Pathologists' recommendations to stain 20 positive and 20 negative tissues to demonstrate accuracy, we recommend the verification of prognostic IHC tests to reduce this number to 15 positive (of which at least 3 weakly positive) and 15 negative tissues without affecting patient safety as illustrated in Table 3 and Figure 1.[11,12,18,19]

For predictive IHC tests, we recommend following the (inter)nationally published guidelines, that is, 20 positive (with at least 5 weakly positive) and 20 negative tissues to demonstrate the accuracy of the tests (Table 3 and Fig. 1).[3,11,16,18]

In case of rare diseases/tumors, the mentioned instructions for type 1 IHC tests (section 2.1.1) can be applied to reach 20 positive and 20 negative specimens. Awaiting full verification/validation, the patient report should include a disclaimer stating that the test was not fully validated and that the results should be interpreted with caution.

Initial Verification of CE-IVD Immunohistochemical Type 1 and 2 Tests for Laboratories With Demonstrable Experience

In view of daily practice, many laboratories have experience with the most common diagnostic (type 1) IHC



tests. To guarantee the balance between efficiency and patient safety, the expert group defined criteria enabling laboratories to demonstrate their experience in an objective way and subsequently apply a simplified procedure (smaller sample size) to verify the IHC biomarker assay. This can be used, for example, when a detection kit of an existing verified CE-IVD IHC test is changed (eg, other vendors, the package insert of the Ab does not contain specifications about the detection kit, and at most about the detection method). The simplified procedure can only be applied when the experience with the method (same principle of detection method, eg, 2-step polymer, 3-step polymer) is demonstrable, the device has not been modified, and/or the same interpretation/scoring method is applied (Fig. 1 and Tables 2 and 3).

Validation of Modified CE-IVD Immunohistochemical Tests With or Without Reference and Non–CE-IVD Immunohistochemical Tests (Laboratory-developed Test)

Within the framework of the IVDR, every change made to a CE-IVD test and every selection of a non–CE-IVD test has to be endorsed by a risk analysis by each individual laboratory to assess the impact on the general safety and the performance characteristics.

Validation of Modified CE-IVD Immunohistochemical Tests With Reference

When, after optimization, the adjustments made to the staining protocol do not meet the manufacturer's

FIGURE 1. Verification of CE-IVD IHC tests according to IFU or validation of modified CE-IVD IHC tests with reference. (1) Type 1 (diagnostic): laboratories should retrospectively and objectively demonstrate their experience for at least 8 of the 15 Ab biomarkers within the verification panel [2 Abs with cytoplasmic staining pattern (C), 2 Abs with nuclear staining pattern (N), 2 Abs with membranous staining pattern (M), and 2 Abs with a more complex staining protocol]. Criteria to demonstrate experience with the type 1 IHC method are, for example, at least good or optimal EQA results for the most recent EQA program within the past 7 years without any change in staining pattern (Ab-antigen interaction) and interpretation of the results (eg, scoring algorithm), traceable IQC results over a certain period (at least 6 mo) or no registration of nonconformities with regard to the method/technique used within the past year. Type 2 (prognostic/pharmaco-predictive): laboratories should demonstrate experience with the test through at least good or optimal EQA results for at least 2 consecutive EQA programs and through concordant inter/intraobserver results. (2) Method verification: a panel of 15 Ab biomarkers is selected for each method (same detection system and staining platform). This panel of 15 Abs, with different pretreatment buffers, consists of 3 Abs with a membranous staining pattern (M), 3 Abs with a cytoplasmic staining pattern (C), 3 Abs with a nuclear staining pattern (N), and 6 Abs with a technical more complex or challenging staining protocol (eg, concentrated Ab). The number of type 1 Abs can be reduced when type 2 Abs using the same detection method and platform have been verified according to the specific type 2 verification requirements (eg, using 3 type 2 and 12 type 1 Abs instead of 15 type 1 Abs). The verification of the accuracy of the method with $5^+/5^-$ or $10^+/10^-$ tissues (negative internal control cells can be present in the tissue sections with antigen-positive cells) is applicable for each Ab 1 to Ab 15. The repeatability/reproducibility is demonstrated by staining at least 3t in triplicate (3s). The repeatability can be demonstrated by staining 3t on at least 3 different positions in the staining platform. Depending on the staining platform, a minimum of 3 slides (3 specimens on 1 slide) and a maximum of 9 slides are distributed within one run. The reproducibility among different runs can be demonstrated by 2 additional stains in 2 other, additional runs. From the 16th Ab marker onwards, a verification of the accuracy with $2^+/2^-$ tissues is recommended after optimization. (3) For the definitions of the performance characteristics, see the section "Definitions" in the main text. (4) The obtained results should be evaluated and compared with the reference or comparator to check whether the predetermined acceptance criteria are met. These criteria are based upon expert consensus results as published in (inter)national publications and differ according to the evaluated performance characteristics and the sample size of the validation set (Tables 2 and 3). Examples are: comparison with (1) the expected results of the reference staining, (2) results of a validated IHC test in the same or in another laboratory, (3) results of the EQA provider, (4) comparator method (eg, ISH test and PCR) and comparison between the results obtained from the within run, between run and readout precision examinations. (5) For a new type 1Ab that does not use the same detection method/staining platform/interpretation, the laboratory checks whether it has experience or not. Subsequently, for each Ab, after optimization, a verification is performed on $5^+/5$ or $10^+/10$ tissues to demonstrate the accuracy and 3t×3s to demonstrate the repeatability/reproducibility. Ab indicates antibody; EQA, external quality assessment; IFU, Instructions For Use; IHC, immunohistochemical; IQ, installation qualification; IQC, internal quality control; ISH, in situ hybridization; OQ, operational qualification; PCR, polymerase chain reaction; PQ, performance qualification; Repeat, repeatability; Reprod, reproducibility; 3s, 3 sections; 3t, 3 tissues.

specifications but a reference for the modified method and its intended use is available, we recommend performing a limited validation in accordance with our instructions for verification of a CE-IVD test according to IFU (Fig. 1 and Tables 2 and 3).

Validation of Modified CE-IVD Immunohistochemical Tests Without Reference

When the modified staining protocol does not meet the manufacturer's specifications and no reference for the modified method and its intended use is available, laboratories should validate the method (eg, modification of pretreatment or detection method) or test (eg, modification to primary Ab) as illustrated in Figure 2 and Tables 2 and 3.

For modified CE-IVD type 2b IHC (predictive) tests with no references available, no instructions for validation were developed as they are considered to be in-house developed tests.

Validation of Non–CE-IVD Immunohistochemical Tests With Reference

The use of non–CE-IVD tests is not recommended but in some cases, there is no suitable alternative available. Examples of non–CE-IVD IHC tests with reference are tests using RUO-marked reagents according to the IFU and non–CE-IVD tests using a (inter)nationally published staining protocol or a nonpublished staining protocol validated by another laboratory. For these tests, we recommend applying the same instructions as for initial validation of modified CE-IVD IHC tests without reference (Fig. 3 and Tables 2 and 3).

Validation of Non–CE-IVD Immunohistochemical Tests Without Reference

Examples of non–CE-IVD IHC tests without reference (not according to IFU) are: (1) tests using RUO-marked reagents with adjustments to the staining protocol, (2) tests using another detection kit (eg other vendors), (3) tests performed on a platform of another manufacturer or non–CE-marked platform, and (4) in house developed tests. These tests require extensive validation as illustrated in Figure 3 and Tables 2 and 3.[3–7,16,18] Various reference techniques are available for this purpose: western blot, enzyme-linked immunosorbant assay, cross-reactivity tests, Ab absorption tests, tests with reagent controls (eg, no primary Ab and isotope), and epitope mapping.[7,27]

Evaluation of Results

The results of the verification/validation examinations should be evaluated and compared with the reference or comparator to check whether predetermined acceptance criteria are met. These criteria are based upon expert consensus results as published in (inter)national publications and differ according to the evaluated performance characteristics and the sample size of the validation set (Tables 2 and 3).[3,6,12]

To verify the accuracy, analytical sensitivity, and analytical specificity, the obtained staining results should be compared with (1) the expected results of the reference staining (eg, correlation with staining characteristics and patterns as described in literature or books), (2) results of a validated IHC test in the same or in another laboratory, (3) results of the EQA provider, or (4) comparator method (eg, in situ hybridization test, polymerase chain reaction).

Staining quality can be objectively evaluated using a scoring system, for example, as introduced by Maxwell and McCluggage[28] (Supplemental Table, Supplemental Digital Content 2, http://links.lww.com/AIMM/A428 scoring criteria for the evaluation of the IHC staining). The parameters of staining intensity and uniformity evaluate the analytical sensitivity of the staining, whereas the absence of background staining evaluates the analytical specificity as well.[28] In addition, the diagnostic sensitivity and specificity of the test can be calculated based on the number of true and false-positive/negative results (Supplemental Table, Supplemental Digital Content 3, http://links.lww.com/AIMM/A429 calculation of diagnostic sensitivity, specificity, and accuracy).[22] This evaluation system can be used during optimization and initial/ongoing (re)verification/validation. For semiquantitative, predictive IHC tests specific algorithms as used in daily practice should be applied, possibly in addition to the mentioned universal scoring system, to calculate the overall concordance between the expected and obtained results.[10,12]

To verify the repeatability/reproducibility, one may evaluate whether the same results are obtained for each repeated sample or tissue within a run and among different runs. Objective acceptance criteria may consist of a minimum percentage of concordant results.

The results from all verification/validation examinations should meet the predefined acceptance criteria before implementation of the new IHC test in daily practice. If acceptance criteria are not met, a root cause investigation and additional testing should be performed. When the results still do not meet the predefined acceptance criteria, the laboratory director should decide and motivate whether or not a test can be implemented into daily practice. If a not fully validated test is implemented, a disclaimer is included in the patient report stating that the test was not fully validated and the results should be interpreted with caution.

ONGOING VALIDATION

Although initial verification/validation is important, it only documents the test performance at a particular point in time. Therefore, verification/validation is an ongoing process. It should, as far as possible, cover the whole IHC staining process by actively monitoring the critical steps from the pre to postanalytical phase [eg, sample processing, equipment maintenance and repair, reagent control/acceptance testing, internal quality control (IQC), training, readout precision control, EQA, concordance studies with non-IHC methods, and population studies].

Internal Quality Control

Daily quality control using internal/external quality control tissue is the best way to monitor and secure staining quality.[11,23,29] Both for optimization, verification/validation, and daily quality control, appropriate control tissue should be selected taking into account the standardization guidelines published by Torlakovic et al.[23,29,30] For diagnostic and prognostic IHC tests, internal normal control tissue (positive and negative internal control cells ideally with different expression levels) present in the patient specimen may be evaluated.[23,29] If no internal control tissue is present and for each predictive IHC test, external control tissue(s) should be placed on the same slide as the patient specimen to allow direct comparison.[23,29] These on-slide IQCs should ideally contain tissue with no, low (close to the detection threshold), intermediate, and high levels of expression of the epitope.[11,16,23,29,31,32] Multitissue blocks can be created for different IHC stainings. These quality control tissues are referred to as "Immunohistochemistry Critical Assay Performance Controls or iCAPs" and enable daily monitoring of the analytical sensitivity and specificity.[11,29,30] In addition, the use of daily internal and on-slide controls continually demonstrates and monitors the repeatability/reproducibility of the IHC test.[11,12]

If control tissue sections are precut and temporarily stored before execution of the IHC test, laboratories should determine the antigen stability for either type 1 and/or type 2 IHC biomarkers within the framework of the robustness of the test as epitope antigenicity gradually drops over time.

External Quality Assurance

Participation in EQA programs allows each laboratory to verify the performance of the IHC test by comparing the obtained results with the "gold standard" results of the EQA provider.[12] A key criterion to define the frequency of participation in EQA programs is the intended use of the IHC test as part of a test panel and interpreted in the context of other morphologic and clinical data. Predictive IHC tests should be assessed at least annually, preferably semiannually.[16,26,28] For prognostic IHC tests, we recommend an EQA participation of at least once every 5 years. For type 1 IHC tests, participation once every 5 years of a selected IHC maker as part of a diagnostic test panel (same detection method) should be sufficient. Risk analysis is also an efficient tool to determine the frequency of EQA participation. If no EQA programs are available alternative approaches should be established, for example, interlaboratory or peer review

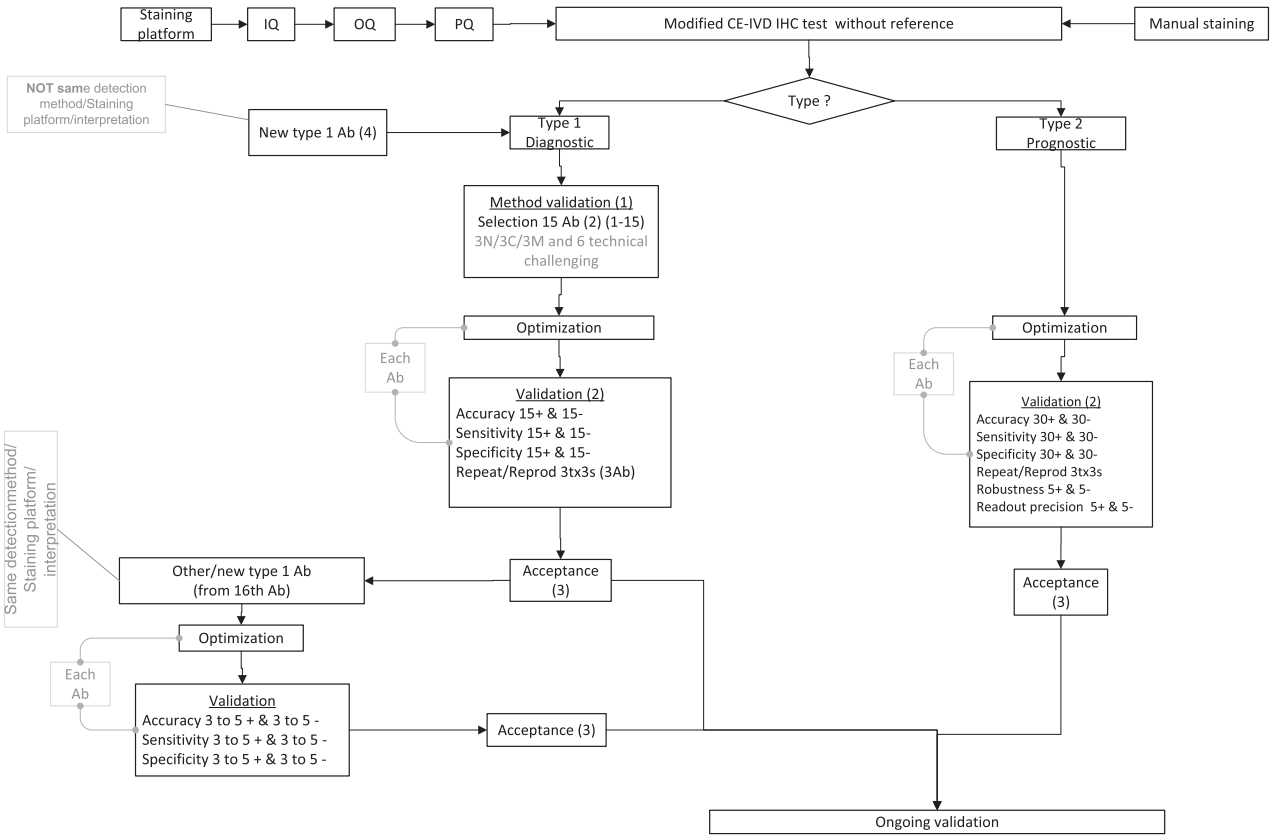


FIGURE 2. Validation of modified CE-IVD IHC tests without reference. (1) Method validation: a panel of 15 Ab biomarkers is selected for each method (same detection system and staining platform). This panel of 15 Abs, with different pretreatment buffers, consists of 3 Abs with a membranous staining pattern (M), 3 Abs with a cytoplasmic staining pattern (C), 3 Abs with a nuclear staining pattern (N), and 6 Abs with a technical more complex or challenging staining protocol (eg, concentrated Ab). The validation of the method with $15^{+}/15^{-}$ tissues (negative internal control cells can be present in the tissue sections with antigen-positive cells) is applicable for each Ab 1 to Ab 15. From the 16th Ab marker onwards, a verification with 3 to 5 positive and 3 to 5 negative tissues (depending on the sample availability) is recommended after optimization. (2) For the definitions of the performance characteristics, see the section "Definitions" in the main text. (3) The obtained results should be evaluated and compared with the reference or comparator to check whether the predetermined acceptance criteria are met. These criteria are based upon expert consensus results as published in (inter)national publications and differ according to the evaluated performance characteristics and the sample size of the validation set (Tables 2 and 3). Examples are: comparison with (1) the expected results of the reference staining, (2) results of a validated IHC test in the same or in another laboratory, (3) results of the EQA provider, (4) comparator method (eg, ISH test, PCR) and comparison between the results obtained from the within run, between run and inter/intraobserver examinations. (4) For a new type 1Ab that does not use the same detection method/staining platform/interpretation, after optimization, a verification is performed on $15^{+}/15^{-}$ tissues to demonstrate the accuracy and 3t×3s to demonstrate the repeatability/reproducibility. Ab indicates antibody; EQA, external quality assessment; IHC, immunohistochemical; IQ, installation qualification; ISH, in situ hybridization; OQ, operational qualification; PCR, polymerase chain reaction; PQ, performance qualification; Repeat, repeatability; Reprod, reproducibility; 3s, 3 sections; 3t, 3 tissues.

testing, reanalyses of previously (before participation at the last EQA program with good or optimal results) examined samples, analyses of registered IQC results, or use of certified (commercially available) reference materials.

CHANGE CONTROL—REVALIDATION

Any modification to the IHC process or staining protocol with an impact on the staining quality and its subsequent results requires revalidation.[3,23] Revalidation may be a result of, for example, deviating EQA results, failing IQC, modification of a lot of a critical reagent, equipment maintenance or repair, modified guidelines, or modified IFU. Modifications can occur at every stage of the IHC process (preanalytic, analytic, and postanalytic; Table 4). Where appropriate, a reoptimization should be performed by applying the mentioned instructions, followed by a revalidation using a limited validation set (Table 4). The extent of the revalidation depends on the impact of the method versus the stand-alone IHC test and its intended use (Table 4) and can be endorsed by a risk analysis.

When the Ab clone (or polyclonal reference) is changed, laboratories should perform a full revalidation equivalent to an initial verification/validation, as a different

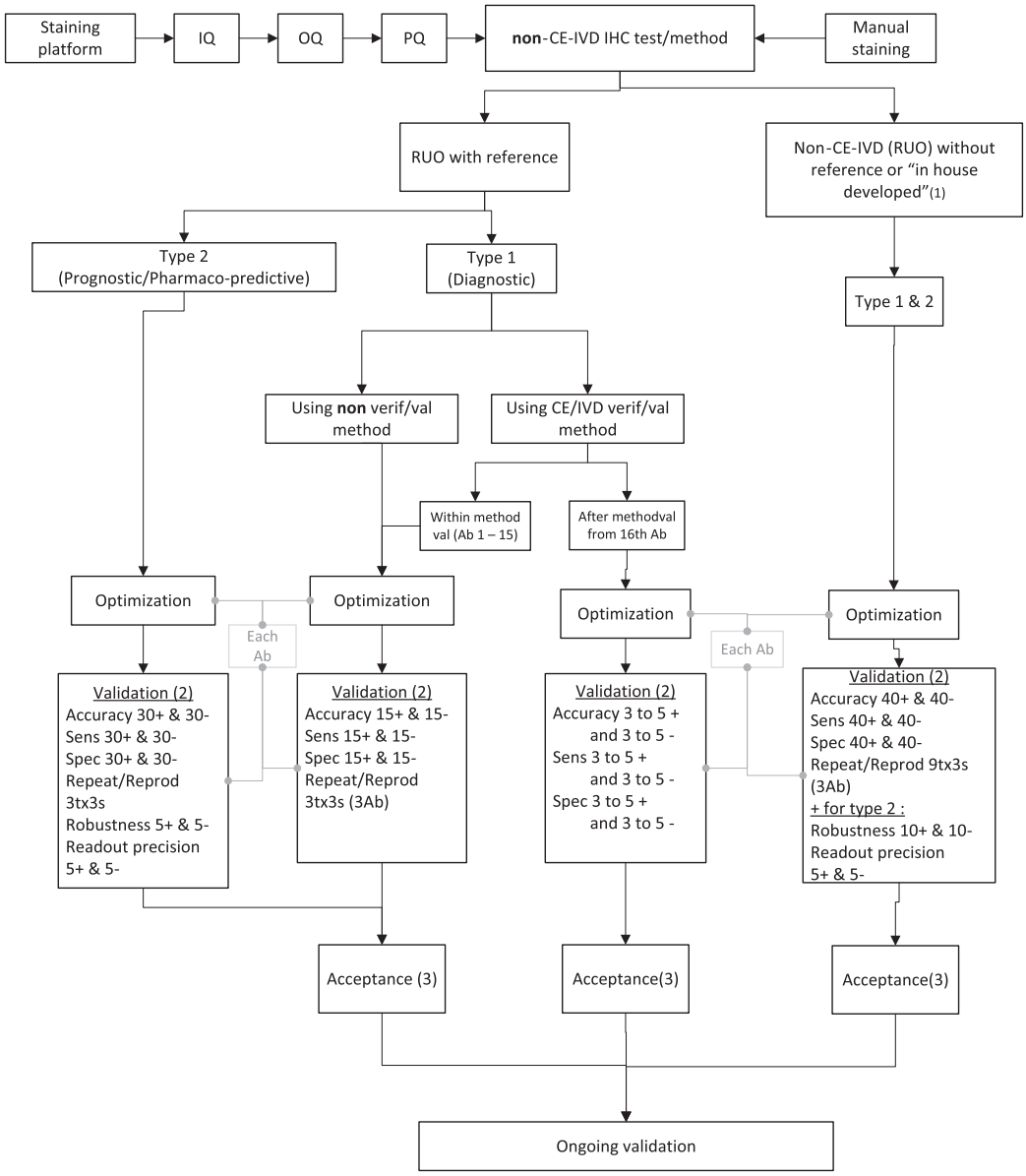


FIGURE 3. Validation of non–CE-IVD IHC tests. (1) For non–CE-IVD type 1 tests without reference and for which limited (control) materials are available, but used within the context of a verified/validated CE-IVD detection method, we recommend starting the initial validation with a set of 2 positive and 2 negative tissues and progressively increase the validation set over time until the target of 40 positive and 40 negative tissues is reached. (2) For the definitions of the performance characteristics, see the section "Definitions" in the main text. (3) The obtained results should be evaluated and compared with the reference or comparator to check whether the predetermined acceptance criteria are met. These criteria are based upon expert consensus results as published in (inter)national publications and differ according to the evaluated performance characteristics and the sample size of the validation set (Tables 2 and 3). Examples are: comparison with (1) the expected results of the reference staining, (2) results of a validated IHC test in the same or in another laboratory, (3) results of the EQA provider, (4) comparator method (eg, ISH test and PCR) and comparison between the results obtained from the within run, between run and inter/intraobserver examinations. Ab indicates antibody; EQA, external quality assessment; IHC, immunohistochemical; IQ, installation qualification; ISH, in situ hybridization; OQ, operational qualification; PCR, polymerase chain reaction; PQ, performance qualification; Repeat, repeatability; Reprod, reproducibility; RUO, Research Use Only; 3s, 3 sections; 3t, 3 tissues.

epitope in the target protein may be detected and subsequently the performance characteristics may significantly vary.[3] Also in case of a modification without reference, a revalidation should be performed following the appropriate instructions for initial validation. Furthermore, a risk analysis should justify this choice.

TABLE 4. Composition of the Validation Set and the Acceptance Criteria for Ongoing Validation and Revalidation of IHC Tests

Modification	Impact	Accuracy	Repeatability	Reproducibility	Others: I/U/S/B/C*
New reagent lot existing IHC test	Stand-alone IHC test	—	—	—	—
T×S†	—	IHC type 1: IQC or 1^+ 1^-	NA	NA	1^+ 1^-
AC (%)	—	≥50	—	—	≥X‡
T×S	—	IHC type 2: 2^+ 2^-	NA	NA	2^+ 2^-
AC (%)	—	≥75	—	—	≥X
Antigen retrieval time antigen retrieval method (e.g. pH, buffer) Ab dilution Ab incubation time vendor (same clone) incubation time detection method	Stand-alone IHC test	—	—	—	—
T×S	—	IHC type 1: 2^+ 2^-	NA	NA	IHC type 1: 2^+ 2^-
AC (%)	—	≥75	—	—	≥X
T×S	—	IHC type 2: 5^+ 5^-	NA	NA	IHC type 2: 5^+ 5^-
AC (%)	—	≥90	—	—	≥X
Type of fixative tissue processing type of paraffin (melting temperature, composition) water supply Ab diluent	Method	—	—	—	—
T×S	—	Panel type 1 IHC: N: 2^+ and 2^- C: 2^+ and 2^- M: 2^+ and 2^- + Panel type 2 IHC: N: 5^+ and 5^- C: 5^+ and 5^- M: 5^+ and 5^-	NA	NA	Panel type 1 IHC: N: 2^+ and 2^- C: 2^+ and 2^- M: 2^+ and 2^- + Panel type 2 IHC: N: 5^+ and 5^- C: 5^+ and 5^- M: 5^+ and 5^-
AC (%)	—	≥87 for type 1 IHC ≥95 for type 1 IHC	—	—	≥X
Equipment (same method)	Method	—	—	—	—
T×S	—	Panel type 1 IHC: N: 2^+ and 2^- C: 2^+ and 2^- M: 2^+ and 2^- + Panel type 2 IHC: N: 5^+ and 5^- C 5^+ and 5^- M: 5^+ and 5^-	3×3 (1Ab)	3×3 (1Ab)	Panel type 1 IHC: N: 2^+ and 2^- C: 2^+ and 2^- M: 2^+ and 2^- + Panel type 2 IHC: N: 5^+ and 5^- C 5^+ and 5^- M: 5^+ and 5^-
AC (%)	—	≥87 for type 1 IHC ≥95 for type 1 IHC	≥90	≥90	≥X
Environmental conditions (eg, relocation)	Method	—	—	—	—
T×S	—	Panel type 1 IHC: N: 1^+ and 1^- C: 1^+ and 1^- M: 1^+ and 1^- + Panel type 2 IHC: N: 1^+ and 1^- C 1^+ and 1^- M: 1^+ and 1^-	3×3 (1Ab)	3×3 (1Ab)	Panel type 1 IHC: N: 1^+ and 1− C: 1^+ and 1^- M: 1^+ and 1^- + Panel type 2 IHC: N: 1^+ and 1^- C 1^+ and 1^- M: 1^+ and 1^-
AC (%)	—	≥83	≥90	≥90	≥X

TABLE 4. (*continued*)

Modification	Impact	Accuracy	Repeatability	Reproducibility	Others: I/U/S/B/C*
Equipment maintenance and repair Change information system/equipment software (with impact on staining results)	Method	—	—	—	—
T×S	—	1 Ab type 1: 2^+ 2^- + 1 Ab type 2: 2^+ 2^-	D§	D	1 Ab type 1: 2^+ 2^- + 1 Ab type 2: 2^+ 2^-
AC (%)	—	≥75	—	—	≥X
Scoring algorithm; AC (%)	Readout precision control Revision of 5^+ and 5^- tissues AC: ≥90				
Ab clone	Full revalidation				

*I = intensity, U = uniformity, S = specificity, B = background, C = counterstaining.
†T = number of tissues/samples and S = number of sections/slides of each selected tissue/sample.
‡Predefined minimum total score.
§If necessary, depending on the part that has been replaced. The number of samples and the repeat factor to verify the repeatability depend on the working principle of the equipment.
Ab indicates antibody; AC, acceptance criterion; C, cytoplasmic; IHC, immunohistochemical; IQC, internal quality control; M, membranous staining pattern; N, nuclear; NA, not applicable.

DISCUSSION

The IVDR does not regulate the operation of IVDs, as this has to be done at the member-state level. Despite the clinical importance and the classification of most IHC tests as class C, currently, clear and useful guidelines on standardization for validating and revalidating these tests are lacking. For this reason, Sciensano established unambiguously, achievable, and affordable instructions for the verification and validation of the analytical performance of all types of IHC tests which should be applicable for each laboratory regardless of its size, diversity in the available tests, and expertise.

Neither CE-IVD tests (according to IFU) nor LDTs fully guarantee successful IHC staining. Results from (inter)national EQA show that the success rate depends on the combination of the Ab clone (or polyclonal reference), the detection system, the staining protocol, and the equipment (total process). Hence, both CE-IVD tests according to IFU and LDTs can be used provided that these tests are respectively verified and validated.[33] In addition to verification/validation, in the application of the IVDR, the laboratory will have to provide a justification for the use of an LDT compared with a CE-IVD test according to the IFU. The use of a CE-IVD test according to the IFU should be the first choice, but when initial testing results do not meet the predetermined acceptance criteria (ie, manufacturer's specifications), the use of an LDT can be an equivalent alternative, in the patient's best interest and as a response to fast-evolving scientific developments. Therefore, Belgian recommendations for validating both CE-IVD IHC tests according to the IFU, as well as LDTs, were established.

Article 5.5 of the IVDR provides the possibility of a partial exemption from the requirements for LDTs.[20] Within this framework, there is a large degree of overlap between the requirements set in the ISO 15198:2012–based Belgian Practice Guideline and those listed in Annexure I of the IVDR.[1,20] In addition, by performing a literature study and an extended risk analysis, we established specific instructions for the analytical verification and validation of both CE-IVD tests and LDTs and incorporated them in the Belgian Practice Guideline enabling the use of LDTs. Therefore, LDTs developed and validated in licensed laboratories will comply with Annexure I of the IVDR.

Despite the fact a clear definition of modifications of CE-IVD tests is missing in Article 2 of the IVDR, we clarified that all modifications of a CE-IVD test outside the IFU are classified as LDT. While the Medical Device Coordination Group will categorize LDTs in (1) CE-IVD with nonsignificant modifications, (2) CE-IVD off-label with significant modifications, (3) RUO, and (4) in-house developed tests, we propose to categorize LDTs according to the origin (modified CE-IVD or non–CE-IVD) of the test and the presence or absence of a reference. The distinction between a CE-IVD test with nonsignificant modification and an off-label CE-IVD test with significant modification would be a very difficult task due to a lack of a clear definition. Even when these definitions should be elaborated, it would be very difficult to interpret and apply them because the patient impact has also been taken into consideration. So, it is not the modification itself that determines the extent of the validation but the final risk incurred by the patient. Hence, we believe that a risk-based categorization of LDT will result in a more objective, harmonized, and standardized approach for

validating the test and subsequently demonstrating its analytical performance.

Although the composition of a validation set is an important factor, we recognize that the availability of the appropriate minimum number of control and patient samples, and of the appropriate resources and trained staff, is often a limiting factor. Therefore, the expert group defined achievable sample sizes and took into account the experience a laboratory might have with IHC biomarker assays. Although these recommendations are not based on statistical analysis, it is mainly a practical, pragmatic, and risk-based approach. The application of method verification/validation and the smaller sample size, without compromising test reliability and patient safety, results in a more achievable and affordable verification/validation procedure. If we also take into account the experience of laboratories with IHC methods, the sample size can be further lowered reducing the workload and subsequent costs. This methodology can also be applied to tests already in use in daily practice but not yet verified or validated. In this case, a retrospective, historical verification through revision of archived material can be performed, respecting the criteria for demonstrating experience as defined in this article. Patient safety is secured by the fact that there is an emphasis shift from initial to ongoing validation of which IQC is obligatory in Belgian laboratories for every IHC test, regardless of its intended use. In addition to ongoing validation and revalidation, a well-elaborated QMS and its continuous evaluation for which guidelines are incorporated in the Belgian Practice Guideline will also support and improve the entire analytical process.

CONCLUSION

Our recommendations are intended to help AP laboratories improve, harmonize, and standardize their validation procedure by providing step-by-step instructions and defining the performance characteristics, sample size, and corresponding acceptance criteria for each of the different types of tests and their intended purposes. In addition, our article defines a clearer distinction between CE-IVD according to the IFU and LDT and categorizes LDT according to a risk-based approach. These efforts towards standardization would very likely enhance the compliance of the IVDR without compromising general safety, increase the range of validated IHC tests, and reduce the cost of validation studies and the associated workload. We believe that the application of these recommendations, together with participation in EQA programs will improve the accuracy of IHC testing, reduce interlaboratory variation, and finally increase the overall quality.

REFERENCES

1. Belgian Commission of Anatomic Pathology. Practice Guideline for setting up a quality system in laboratories for anatomic pathology in the context of the licensing decree. Accessed November 7, 2022: https://www.sciensano.be/fr/biblio/directive-pratique-anatomie-pathologique
2. Verbeke H, Dierick AM, Ghislain V, et al. Analytical validation of tests in laboratories of anatomic pathology: a Belgian population-based study. *Accred Qual Assur*. 2020;25:69–79.
3. Fitzgibbons PL, Bradley LA, Fatheree LA, et al. Principles of analytic validation of immunohistochemical assays: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch Pathol Lab Med*. 2014;138:1432–1443.
4. Burd EM. Validation of laboratory-developed molecular assays for infectious diseases. *Clin Microbiol Rev*. 2010;23:550–576.
5. Rabenau HF, Kessler HH, Kortenbusch M, et al. Verification and validation of diagnostic laboratory tests in clinical virology. *J Clin Virol*. 2007;40:93–98.
6. Mattocks CJ, Morris MA, Matthijs G, et al. A standardized framework for the validation and verification of clinical molecular genetic tests. *Eur J Hum Genet*. 2010;18:1276–1288.
7. Smith NR, Womack C. A matrix approach to guide IHC-based tissue biomarker development in oncology drug discovery. *J Pathol*. 2014;232:190–198.
8. Raymaekers M, Smets R, Maes B, et al. Checklist for optimization and validation of real-time PCR assays. *J Clin Lab Anal*. 2009;23:145–151.
9. Elliott K, McQuaid S, Salto-Tellez M, et al. Immunohistochemistry should undergo robust validation equivalent to that of molecular diagnostics. *J Clin Pathol*. 2015;68:766–770.
10. Cheung CC, D'Arrigo C, Dietel M, et al. Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine: part 1: fit-for-purpose approach to classification of clinical immunohistochemistry biomarkers. *Appl Immunohistochem Mol Morphol*. 2017;25:4–11.
11. Cheung CC, D'Arrigo C, Dietel M, et al. Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine: part 4: tissue tools for quality assurance in immunohistochemistry. *Appl Immunohistochem Mol Morphol*. 2017;25:227–230.
12. Torlakovic EE, Cheung CC, D'Arrigo C, et al. Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine. part 3: technical validation of immunohistochemistry (IHC) assays in clinical IHC laboratories. *Appl Immunohistochem Mol Morphol*. 2017;25:151–159.
13. Torlakovic EE, Cheung CC, D'Arrigo C, et al. Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine—part 2: immunohistochemistry test performance characteristics. *Appl Immunohistochem Mol Morphol*. 2017;25:79–85.
14. Ardelean AI, Catoi C. The validation of routine analytical methods in histotechnology: a practical approach. *Vet Med*. 2011;1:42–50.
15. International Organization of Standardization. *ISO 15189 Medical Laboratories - Requirements for quality and competence*, 3rd edn. ISO (International Organization of Standardization); 2012.
16. Lambein K, Guiot Y, Galant C, et al. Update of the Belgian guidelines for HER2 testing in breast cancer. *Belg J Med Oncol*. 2014;8:109–115.
17. Jouret-Mourin A, Hoorens A, Kockx M, et al. Belgian guidelines for HER2 testing in gastric cancer. *Belg J Med Oncol*. 2011;5:14–22.
18. Fitzgibbons PL, Murphy DA, Hammond MEH, et al. Recommendations for validating estrogen and progesterone receptor immunohistochemistry assays. *Arch Pathol Lab Med*. 2010;134:930–935.
19. Hammond MEH, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol*. 2010;28:2784–2795.
20. European Commission. Regulation (EU) 2017/746 of the European Parliament and of the council_of 5 April 2017 on in vitro diagnostic medical devices and repealing_directive 98/79/EC and commission decision 2010/227/EU. May 5, 2017. https://eur-lex.europa.eu/eli/reg/2017/746/oj
21. Jennings LJ, Arcila ME, Corless C, et al. Guidelines for validation of next-generation sequencing–based oncology panels: a Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn*. 2017;19:341–365.

22. Baratloo A, Hosseini M, Negida A, et al. Part 1: simple definition and calculation of accuracy, sensitivity and specificity. *Emerg (Tehran)*. 2015;3:48–49.
23. Torlakovic EE, Riddell R, Banerjee D, et al. Canadian Association of Pathologists–Association canadienne des pathologistes National Standards Committee/Immunohistochemistry: best practice recommendations for standardization of immunohistochemistry tests. *Am J Clin Pathol*. 2010;133:354–365.
24. Satturwar S, Malenie R, Sutton A, et al. Validation of immunohistochemical tests performed on cytology cell block material: practical application of the College of American Pathologists guidelines. *Cytojournal*. 2019;16:6.
25. Stuart LN, Volmar KE, Nowak JA, et al. Analytic validation of immunohistochemistry assays: new benchmark data from a survey of 1085 laboratories. *Arch Pathol Lab Med*. 2017;141:1255–1261.
26. Fitzgibbons PL, Goldsmith JD, Souers RJ, et al. Analytic validation of immunohistochemical assays. *Arch Pathol Lab Med*. 2017;141: 1247–1254.
27. Bordeaux J, Welsh AW, Agarwal S, et al. Antibody validation. *Biotechniques*. 2010;48:197–209.
28. Maxwell P, McCluggage W. Audit and internal quality control in immunohistochemistry. *J Clin Pathol*. 2000;53:929–932.
29. Torlakovic EE, Nielsen S, Francis G, et al. Standardization of positive controls in diagnostic immunohistochemistry: recommendations from the International Ad Hoc Expert Committee. *Appl Immunohistochem Mol Morphol*. 2015;23:1–18.
30. Torlakovic EE, Francis G, Garratt J, et al. Standardization of negative controls in diagnostic immunohistochemistry: recommendations from the international ad hoc expert panel. *Appl Immunohistochem Mol Morphol*. 2014;22:241–252.
31. Allison KH, Hammond MEH, Dowsett M, et al. Estrogen and progesterone receptor testing in breast cancer: ASCO/CAP guideline update. *J Clin Oncol*. 2020;38:1346–1366.
32. Wolff AC, Hammond MEH, Allison KH, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline focused update. *J Clin Oncol*. 2018;36:2105–2122.
33. Kaul KL, Sabatini LM, Tsongalis GJ, et al. The case for laboratory developed procedures: quality and positive impact on patient care. *Acad Pathol*. 2017;4:2374289517708309.