

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/173020>

Please be advised that this information was generated on 2019-04-20 and may be subject to change.

## **Article 25fa pilot End User Agreement**

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with explicit consent by the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU) 'Article 25fa implementation' pilot project. In this pilot research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and/or copyrights owner(s) of this work. Any use of the publication other than authorised under this licence or copyright law is prohibited.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please contact the Library through email: [copyright@ubn.ru.nl](mailto:copyright@ubn.ru.nl), or send a letter to:

University Library  
Radboud University  
Copyright Information Point  
PO Box 9100  
6500 HA Nijmegen

You will be contacted as soon as possible.

# 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

Carol C. Cheung, MD, PhD, JD,\*† Corrado D'Arrigo, MB, ChB, PhD, FRCPath,‡§||  
 Manfred Dietel, MD, PhD,¶|| Glenn D. Francis, MBBS, FRCPA, MBA, FFSc (RCPA),#\*\*††  
 C. Blake Gilks, MD,‡‡ Jacqueline A. Hall, PhD,§§|| Jason L. Hornick, MD, PhD,¶¶||  
 Merdol Ibrahim, PhD,## Antonio Marchetti, MD, PhD,\*\*\* Keith Miller, FIBMS,##  
 J. Han van Krieken, MD, PhD,††† Soren Nielsen, BMS,‡‡‡§§§ Paul E. Swanson, MD,||| ||  
 Clive R. Taylor, MD,¶¶¶ Mogens Vyberg, MD,‡‡‡§§§ Xiaoge Zhou, MD,###\*\*\*\*  
 and Emina E. Torlakovic, MD, PhD,\*††††††††

From the International Society for Immunohistochemistry and Molecular Morphology (ISIMM)  
 and International Quality Network for Pathology (IQN Path)

**Abstract:** Technical progress in immunohistochemistry (IHC) as well as the increased utility of IHC for biomarker testing in precision medicine avails us of the opportunity to reassess clinical IHC as a laboratory test and its proper characterization as a special type of immunoassay. IHC, as used in current clinical applications, is a descriptive, qualitative, cell-based, usually nonlinear, in situ protein immunoassay, for which the readout of the results is principally performed by pathologists rather than by the instruments on which the immunoassay is performed. This modus operandi is in contrast to other assays where the instrument also performs the readout of the test result (eg, nephelometry readers, mass spectrometry readers, etc.). The readouts (results) of IHC tests are used either by pathologists for diagnostic purposes or by treating physicians (eg, oncologists) for patient management decisions, the need for further testing, or follow-up. This paper highlights the distinction between the

original purpose for which an IHC test is developed and its subsequent clinical uses, as well as the role of pathologists in the analytical and postanalytical phases of IHC testing. This paper is the first of a 4-part series, under the general title of “Evolution of Quality Assurance for Clinical Immunohistochemistry in the Era of Precision Medicine.”

**Key Words:** biomarkers, quality assurance, quality control, validation, immunohistochemistry

(*Appl Immunohistochem Mol Morphol* 2017;25:4–11)

In the era of precision medicine, biomarker testing using immunohistochemistry (IHC) has not only become more precise but also more complex.<sup>1–6</sup> Precision medicine requires precision results, which can only come about from precision testing. Because of increasing reliance on

Received for publication August 1, 2016; accepted August 26, 2016.

From the \*Department of Laboratory Medicine and Pathobiology, University of Toronto; †Department of Pathology, University Health Network; †††Department of Laboratory Hematology, University Health Network, Toronto, ON; ‡‡Vancouver General Hospital, University of British Columbia, Vancouver, BC; ||||Department of Pathology and Laboratory Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada; ‡Poundbury Cancer Institute; §Dorset County Hospital NHS Foundation Trust; ||Cancer Diagnostic Quality Assurance Services (CADQAScic), Dorchester; |||Division of Cancer, Department of Surgery & Cancer, Imperial College, London, UK; ##UK National External Quality Assessment Scheme (UK NEQAS), University College London, London, UK; ¶|School of Medicine, Institute of Pathology, Charité—University Hospital Berlin, Berlin, Germany; #Griffith University, Gold Coast; \*\*RCPA Quality Assurance Program, Sydney; ††Genomics For Life, Brisbane, Australia; §§International Quality Network for Pathology (IQN Path), Luxembourg City, Luxembourg; ¶¶Brigham and Women's Hospital, Harvard Medical School, Boston, MA; \*\*\*Center of Predictive Molecular Medicine, Center for Excellence on Ageing and Translational Medicine, University of Chieti-Pescara, Chieti, Italy; †††Department of Pathology, Radboud University Medical Centre, Nijmegen, The Netherlands; ‡‡‡Institute of Pathology, Aalborg University Hospital and Department of Clinical Medicine, Aalborg University; §§§Nordic Immunohistochemistry Quality Control (NordiqC), Aalborg, Denmark; ¶¶¶Keck School of Medicine, University of Southern California, Los Angeles, CA; ###Department of Pathology, Beijing Friendship Hospital, Capital Medical University, Beijing, China; \*\*\*\*Chinese Committee for Pathologists—Immunohistochemistry Quality Control; and ††††Canadian Immunohistochemistry Quality Control (CIQC)/Canadian Association of Pathologists National Standards Committee for High Complexity Testing/Immunohistochemistry, Vancouver, BC, Canada.

The authors declare no conflict of interest.

Reprints: Emina E. Torlakovic, MD, PhD, Department of Laboratory Hematology, Toronto General Hospital/UHN, 200 Elizabeth Street, Toronto, ON, Canada M5G 2C4 (e-mail: emina.torlakovic@uhn.ca).

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

the results of biomarker testing in patient management, a clear understanding by the pathologist of the tests themselves that provide this information is critical to appropriate patient care and thus patient safety.<sup>7-9</sup>

In this paper, we emphasize the essential role of “purpose” in the IHC assay:

- that “purpose” is the intended use of an IHC test at the time that the test was developed,
- that “purpose” is intrinsic to the identity of any particular IHC test, and
- that classification based on “purpose” of the IHC test is helpful in developing proper quality assurance tools and clarifies the utilities of that test in patient management in the era of precision medicine.<sup>10</sup>

### IHC AS A QUALITATIVE ASSAY FOR THE IN SITU DETECTION OF PROTEIN BIOMARKERS

It is essential to recognize that in its most common application, IHC is a descriptive, threshold-based test. In this context it is generally unknown to what extent the relationship between the amount of target protein and observed intensity of achieved signal is linear.<sup>11-14</sup> The intention is to deliver a “positive” or “negative” signal where appropriate and meaningful, rather than to measure the amount of target protein. This mode of application is similar to other testing methods that employ amplification such as polymerase chain reaction (PCR)-based methods, which are also considered as descriptive or “threshold methodologies.” For example, the intention of performing PCR testing is to demonstrate whether DNA or RNA sequences of interest are detected or not, while in IHC the focus is on whether a target protein-based biomarker is detected or not; however, both need to be demonstrated at clinically relevant sensitivity and specificity, which is achieved through validation.<sup>15</sup>

Today the results of IHC testing are increasingly applied as biomarkers. The term “biomarker” (or biological marker) is broadly defined as any biological or physiological moiety that is used to identify disease, guide targeted therapy or monitor for recurrence.<sup>16,17</sup> Although the nomenclature is not fully standardized, the National Institutes of Health working group’s definition of a biomarker is more rigid and requires that a biomarker can be objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic response to a therapeutic intervention.<sup>18</sup> The American Association of Pharmaceutical Scientists and the US Clinical Ligand Society have identified 4 general classes of biomarker assays: (i) definite quantitative assays (eg, mass spectrometry; reference standard is well defined and fully representative of the endogenous biomarker); (ii) relative quantitative assays (eg, ligand-binding assay; reference standards may not be available in a defined form or fully representative of an endogenous biomarker); (iii) quasiquantitative assays (eg, qRT-PCR, which does not employ the use of a calibration standard, but has a continuous response and the result is expressed in terms of a characteristic of the test sample, and (iv) qual-

itative assays (eg, PCR or IHC; these tests generate categorical data that lack proportionality to the amount of analyte in a sample.<sup>8</sup> Qualitative assays such as PCR or IHC typically generate nominal results (eg, presence or absence of target analyte; eg, PCR test positive for *Mycobacterium tuberculosis* or IHC test positive for S-100 or anaplastic lymphoma kinase (ALK)).<sup>8</sup> The results of qualitative assays may also be expressed in ordinal form by use of scoring scales for readout (eg, 0 to 3+ for HER2 IHC in breast cancer). Some IHC tests may have complex readout rules and require either percentage estimate or direct cell counting with predefined cutoff points for what is considered a “positive” or a “negative” result. An example of this approach is the PD-L1 IHC 22C3 PharmDx assay (Dako/Agilent Technologies, Canada) where the cutoff point for a positive test is 50% positive tumor cells.<sup>19</sup> Although the percentage of positive cells is reported, no measurement is involved as the overall result is based on identifying the threshold between a positive cell (any membranous staining of tumor cells) versus a negative cell (no staining of tumor cell membrane). Ultimately, therefore, this biomarker is descriptive and qualitative, rather than quantitative in a strict sense.

The future development of IHC and related methods for protein quantification, in the sense of measurement of in situ protein, is beyond the purview this paper and functional capability of IHC as currently practiced. However, the qualitative nature of IHC, as used today in clinical practice, does not make IHC biomarkers inferior to other biomarkers. IHC biomarkers are useful diagnostic, prognostic, and predictive markers when properly validated for specific use.<sup>20</sup> Applying identical logic, the PCR test is a valid biomarker for detection of *M. tuberculosis*, or B or T cell clonality, if the test is properly validated for the specified use.<sup>21,22</sup> What ultimately makes a laboratory test “bad” or “good” does not necessarily depend on the level of quantitation in methodology nor on the sophistication of the methodology, but rather, on whether or not the test is properly validated (clinically, diagnostically, and technically) for the specified purpose (“fit-for-purpose”).<sup>10</sup>

### “PURPOSE” IS INTRINSIC TO THE IHC TEST

“Purpose” is intrinsic to the identity of every laboratory test, including IHC tests.<sup>8,10,23,24</sup> Because of increasing awareness of the need for quality assurance in every aspect of IHC testing, it is important that the development of each IHC-based assay takes into consideration its intended use at the time of its development. This is particularly relevant to the validation of IHC tests. As we will show in this series, the type of validation depends entirely on the purpose for which a test is developed.<sup>10,15,25-27</sup> When the purpose of a test is changed, even if it identifies the same target molecule, it becomes a different test and therefore its validation requirements should be revisited. This can be illustrated by considering the use of antibodies to CD34. Table 1 describes 3 different purposes for performing a CD34 IHC test in clinical practice. As there are at least 3 different purposes for identifying the

**TABLE 1.** Different Purposes of the CD34 Immunohistochemistry (IHC) Test

IHC Test	Field	Purpose	Readout
CD34	Dermatopathology and soft tissue pathology	To distinguish solitary fibrous tumor (CD34 <sup>+</sup> ) from desmoplastic mesothelioma (CD34 <sup>-</sup> ), or To distinguish dermatofibroma (CD34 <sup>-</sup> ) from dermatofibrosarcoma protuberans (CD34 <sup>+</sup> )	Positive vs. negative
CD34	Hematopathology	To identify CD34 <sup>+</sup> blasts in bone marrow tissue biopsy, to distinguish between acute leukemia (> 20% CD34 <sup>+</sup> blasts) and myelodysplastic syndrome (MDS) as well as aplastic anemia (no or decreased CD34 <sup>+</sup> blasts) from hypoplastic MDS (increased CD34 <sup>+</sup> cells)	Enumeration of CD34 <sup>+</sup> blasts Counting 500 to 1000 cells Cutoff points at 5% (increased blasts), 10% (accelerated phase of myeloproliferative neoplasms or diagnosis of MDS-II), and > 20% (acute leukemia)
CD34	Vascular density	To determine degree of vascularization	IA using microvessel density algorithm

IA indicates image analysis.

CD34 molecule, this in fact represents 3 different CD34 tests, each with its own purpose, test performance characteristics, validation requirements and scoring/readout criteria; this holds true even if for some (different) purposes the CD34 IHC protocol remains the same (Fig. 1).<sup>28-31</sup>

The question of whether there is such a thing as a “multipurpose” IHC test inevitably arises and it is an important question. The answer is that there can be no such thing as a “multipurpose IHC test” for the reason that purpose is intrinsic to the very definition of any laboratory test including all IHC tests. However, an IHC protocol may certainly be multipurpose, where a single set of protocol conditions may be applied for different IHC tests. For example, it is possible that a single IHC protocol can be used for the purpose of detecting PAX8 expression in tumors of thyroid, kidney, and Müllerian origin. However, it is also possible that an IHC protocol designed to detect PAX8 expression in thyroid tumors, where expression levels are usually high, may fail to detect expression of PAX8 in some other tumors with possibly lower expression levels. Therefore, in the development of this hypothetical PAX8 IHC assay, we would need to consider whether the assay’s analytical sensitivity is fit for each of our various purposes. Although in current clinical practice the majority of IHC protocols are performed under the assumption that they are “multipurpose,” it is unlikely that many such protocols were actually validated for all purposes for which they are being used. Data from proficiency testing (NordQC, UK NEQAS, etc.) clearly indicates that validation of “multipurpose” IHC protocol is challenging and not uniformly performed by all laboratories. For example, for PAX8, run 42 2014 in the NordQC program, 25% of the participating laboratories were able to demonstrate PAX8 in the high level expressing sample but produced a false negative result in the sample with low level expression such as clear cell renal cell carcinoma.<sup>32</sup>

It is of utmost importance that should any new purpose arise for which an existing IHC protocol will be used, it should be carefully re-evaluated and the need for additional validation carefully considered.

### “PURPOSE,” “INTENDED USE,” “USE,” AND “FIT-FOR-PURPOSE”

The “purpose” of an IHC test may be defined as the “intended use at the time the test was developed.” However the “purpose” for which a test was developed and validated by the clinical laboratory may not necessary be the same as the “use” that an ordering physician has in mind (eg, a staff pathologist) after the test has been made available for clinical practice; if such situations arise, then “current use(s)” would no longer align with “intended use at the time the test was developed” (ie, original purpose). For example, the ALK IHC test that was developed and validated for the purpose of detecting anaplastic large cell lymphoma has been “used” to detect ALK expression in non-small cell lung cancer. However, initial attempts resulted in unacceptably low sensitivity because the 2 entities harbor different gene rearrangements and have different protein expression levels; as such, it was discovered that the detection of ALK expression in lung cancer may require an IHC protocol of higher sensitivity than that required for ALK expression in lymphoma.<sup>33</sup> In addition, the readout criteria (eg, cutoff values, etc.), which are quite different in lymphomas and non-small cell lung cancer, also must be adapted to reliably predict the clinical relevance.<sup>34-36</sup> Therefore, for IHC tests, “purpose” and “intended use at the time the test was developed” reflect test development, which is within the domain of the laboratory. In contrast, “use” in a more general sense relates to clinical practice, which may be different from the original “intended use” and is within the domain of the practicing physician (eg, pathologist and/or treating physician). The concept of “fit-for-purpose” has evolved to address such possible confusion. “Fit-for-purpose” describes an assay that has been successfully validated for the intended use at the time the assay was developed, combining both laboratory and clinical definitions.<sup>10,37</sup> It is expected that the new biomarkers will be fit-for-purpose as there are high expectations that the biomarkers will improve diagnosis, define disease subsets that may differ in response, define individual variability in the drug’s molecular target, and provide early clues regarding response to therapy.<sup>38,39</sup>

From the perspective of the laboratory, purpose is the reason for which the laboratory sets up the test and is what dictates how the laboratory performs the validation of that test. It is therefore important that prior to the development, calibration, and validation of a new IHC assay, there should be coordination between the laboratory director and the end user(s) to ensure that intended use (ie, purpose, which will drive validation requirements) and actual use will be properly aligned. Whether an assay that has been introduced into clinical practice in this manner will ever be “used” for a different purpose than originally intended is unknown to both the laboratory director and to end user(s). However, should such a scenario arise (eg, a different clinical use of an existing assay), a new “fit-for-purpose” validation process would be required.

Lastly, an important point to note regarding “purpose” is that in relation to clinical IHC testing, “purpose” is not simply the in situ detection of a biological gene product in tissue sections. In order to be meaningful, the purpose of any clinical IHC test must always be accompanied by a medical context, because it is this context that will impact the validation requirements of a particular IHC test.

### CLASSIFICATION OF IHC TESTS BASED ON FIT-FOR-PURPOSE PRINCIPLES

A number of classification schemes for diagnostic tests have been introduced in different regions of the world as part of regulatory oversight to ensure appropriate testing for patient care.

The main principle underlying the various classification schemes is based on the degree of risk to patient safety and public health. The degree of oversight for a test should be commensurate with the perceived risk to the patient.<sup>40,41</sup> The Food and Drug Administration in the United States (FDA) and International Medical Device Regulators Forum have adopted this principle.<sup>42,43</sup> An approach based on the degree of risk to patient safety translates into the domain of quality assurance, where the test classification dictates differing levels of stringency in relation to requirements for validation, documentation, and reporting. From a quality assurance perspective, approaching the classification of tests in this manner ensures that the “fit-for-purpose” principle is always kept in mind and that the test will be conducted in a manner that considers both the needs of the user and the effect of the test result on the patient.

#### TYPE 1-IHC AND TYPE 2-IHC TESTS

The Canadian Association of Pathologists (CAP-ACP) has proposed 2 classes of IHC tests based on the end user of the IHC results: class I (pathologist end user) and class II (treating physician end user) IHC tests.<sup>44</sup> The current paper adheres to the same principles but offers an update in terminology in order to highlight and clarify the role of IHC test classification in quality assurance. IHC tests where the test is read by the pathologist, and the

**TABLE 2.** The Classification of Immunohistochemistry (IHC) Tests From the Perspective of the End User

Test Category	User-Based Classification	Example
Diagnosis in symptomatic patients (diagnostic)	Pathologist (type 1-IHC)	S100, VIM, CD45 and PAN-CK in diagnosis of unknown primary tumor
Disease screening (for an additional disorder) in symptomatic patients (diagnostic)	Treating physician (type 2-IHC)	MLH1, MSH2, MSH6 and PMS2 in colorectal cancer patients being screened for Lynch syndrome
Prognosis of a diagnosed disease (prognostic)	Treating physician (type 2-IHC)	CD10, Bcl-6, and MUM1 for cell of origin in diffuse large B-cell lymphoma
Predictive of treatment response or adverse reaction (predictive)	Treating physician (type 2-IHC)	ER, PR, HER2 for breast cancer, HER2 for gastric cancer

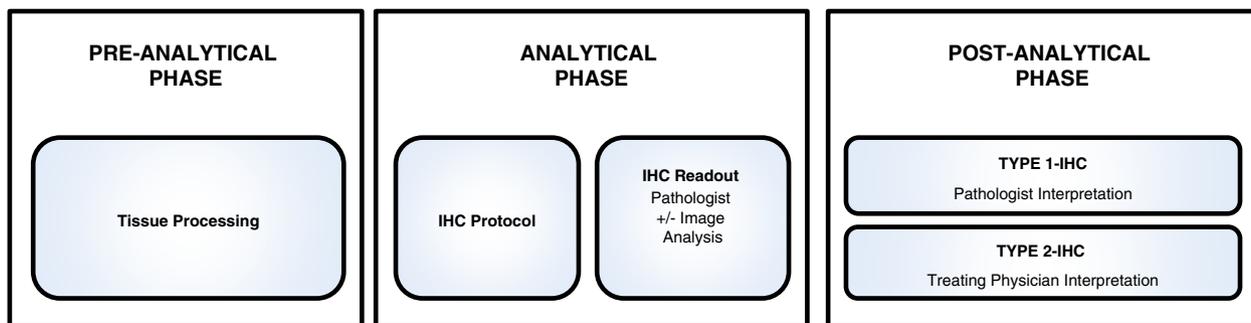
results of the readout used by that pathologist for his/her diagnostic practice are termed “pathologist end user IHC tests” (type 1-IHC tests) (Table 2).<sup>45</sup> IHC tests where the pathologist’s readouts are used by a treating physician in order to determine patient management, are “treating physician end user IHC tests” (type 2-IHC tests). The latter include prognostic tests, predictive tests, and screening tests (a special subtype of diagnostic tests) that are relevant for patient management (Table 2).

Figure 1 illustrates similarities and differences between type 1 and type 2 tests.

Although these definitions of type 1-IHC and type 2-IHC tests have broad correlation to various regulatory frameworks, this terminology is introduced to avoid the use of the term “class,” which has been defined on a different basis. The goal is to avoid potential confusion with FDA classification of IHC devices and other test classification schemes in various countries that address risk and refer to industry and devices rather than IHC tests and how these are actually used by pathologists and/or treating physicians. The comparison with classification schemes used by various regulatory agencies is shown in Table 3.<sup>41,42,44,46,47</sup>

### THE IHC PROTOCOL, THE “READOUT” AND THE “INTERPRETATION”

The IHC test encompasses preanalytic, analytic, and postanalytic phases. Figure 1 illustrates these different components of the IHC test. The analytical phase of the IHC test is complex and it consists of various components of the IHC protocol as well as the “readout” of the generated IHC slide. There is an essential difference between IHC testing and other biomarker testing methods. As noted above, IHC tests are for the most part descriptive, or at best quasiquantitative, and the readout is the culmination of a morphologic analysis of an IHC slide by a pathologist, who assesses the presence and localization of signals, evaluates cutoff points between what



**FIGURE 1.** The phases of IHC testing and the dual role of the pathologist in the analytical and postanalytical phases. The pathologist performs the readout (with or without assistance by image analysis) and is also the end user for type 1-IHC testing. IHC indicates immunohistochemistry.

constitutes a positive and negative signal, and counts positive cells, or uses other methods of scoring as appropriate.<sup>48,49</sup> In essence, the IHC test is always descriptive, even when it is possible to provide information in a quantitative format.

It takes years of training and experience as well as focused continuing medical education to become proficient at reading IHC slides. For certain biomarkers, the readout criteria to determine whether a test is “positive” or “negative” may be seemingly straight forward (eg, reading thyroid transcription factor-1 IHC slides in lung cancer); however, it may also be very complex such as when multiple criteria are incorporated in the readout (eg, reading HER2 IHC slides) or when evaluating for vascular invasion, microinvasion, micrometastasis, where higher levels of experience are required for proper readout.<sup>50–54</sup> This practice is clearly in contrast to other biomarkers for which readouts are performed by automated instruments of different types (nephelometry, colorimetric methods, mass spectrometry, flow cytometry, etc.). Computer aided digital image analysis (IA) has been introduced for readout of IHC slides, with the intent of improving reproducibility of results.<sup>55–58</sup> IA currently still

requires time from a pathologist to determine what part of the tissue and what cells will be scored as well as which algorithms are applicable for a given readout; a number of these algorithms have been approved by the FDA and other organizations for clinical use.<sup>59</sup> In addition, the approval of whole slide imaging for primary diagnosis, already effected in Europe and Canada, and under active trial in the United States, surely will impact the growth of computer-assisted morphometry and analysis of IHC in the near future. Nevertheless, at present, IA should be viewed as assistive technology that helps, but currently cannot replace, pathologists in generating the readout from an IHC slide.<sup>60</sup> The analytical phase of IHC testing is completed by the readout of the IHC slide.

Interpretation of the readout occurs in the post-analytical phase of IHC testing.<sup>61</sup> For type 1-IHC tests, it is the pathologist end user who interprets the readout results. An example of this would be a clinical case of “tumor of unknown origin” where IHC testing was ordered to help identify the potential primary site of a metastatic tumor deposit. A typical readout result might be that cytokeratin 20 and CDX2 are positive and cytokeratin 7 and thyroid transcription factor-1 are negative;

**TABLE 3.** Comparison of Immunohistochemistry (IHC) Classification Schemes of IHC Tests by Regulatory Agency and Classification Based on Clinical Practice

Regulation of Test Manufacturers	Intention	Class	Clinical Practice (Nearest Correlate)
FDA	For regulating manufacturers of tests	Class 1 Class 2 Class 3	Type 1-IHC Type 2-IHC Type 2-IHC
Health Canada	For regulating manufacturers of tests	Class 1 Class 2 Class 3 Class 4	Type 1-IHC Type 2-IHC Type 2-IHC Type 2-IHC
EU In-vitro Diagnostic Regulation	For regulating manufacturers of tests	Class A Class B Class C Class D	NA NA Type 1-IHC, type 2-IHC Type 2-IHC
Canadian Association of Pathologists	Guidance for clinical practice	Class 1 Class 2	Type 1-IHC Type 2-IHC

FDA indicates Food and Drug Administration in the United States; NA, not available.

the subsequent interpretation by the pathologist end user (ie, the sign-out pathologist) may then reasonably be “consistent with adenocarcinoma with intestinal differentiation.”<sup>62</sup> The sign-out pathologist in this example may or may not choose to report the detailed readout of the IHC results, but will certainly report his/her interpretation of the readout results. It is also worth noting that for type 1-IHC tests, which by current convention are both “read” and “interpreted” by the same pathologist (ie, the sign-out pathologist), it may appear as though readout and interpretation are concurrent; in practice they may occur very close together, but the interpretation is necessarily subsequent to the readout.

For type 2-IHC tests, it is the treating physician end user who interprets the readout results in the context of clinical actions. An example would be a treating physician who receives IHC results for a patient with carcinoma of the breast stating that the sampled tumor cells are “strongly positive for estrogen receptor and progesterone receptor and negative for HER2”; the treating physician may then interpret the readout results as indicating that “the patient is a good candidate for hormonal therapy.”<sup>63</sup>

When a pathologist reports a test result as “positive” versus “negative,” the presumption is that the pathologist is following predetermined scoring criteria using predetermined scoring or cut off points, extending beyond the usual meaning of “interpretation.” As the term “interpretation” is inherently associated with post-analytical phase, it is important that any readout as a part of analytical phase, irrespective of its complexity, is not called “interpretation.” Moreover, in some type 2 tests, pathologists may also contribute more directly to the clinical interpretation that will eventually help the treating physician in patient management. For example in IHC testing for DNA mismatch repair deficiency, pathologists may provide additional interpretation of unusual patterns (eg, when reporting discordant subclonal patterns) that directly impact upon how the treating physician interprets and applies the results in treatment.<sup>64</sup>

This paper seeks to apply the standardized terminology for laboratory testing in the context of IHC; thus “readout” is considered to be part of the analytic phase, whereas “interpretation” is linked to the postanalytical phase.<sup>65</sup> For clarity, the term “readout” is applied to the process whereby a pathologist using high-level skills “reads” the IHC slide (ie, produces a result), before a closely related but subsequent process in which the readout (or IHC result) is interpreted in the context of pathology diagnosis or clinical scenario. Clearly, pathologists are “interpreting” (in the more general use of the word) all visual cues present in the IHC slides in arriving at the most accurate readout (result) (eg, assessing signal intensity, localization, “texture,” and other signal parameters in conjunction with controls, excluding various type of artifacts, etc.). However, all subsequent mental processing of the information incorporated into the “readout” is, by analogy with other forms of laboratory testing, properly considered to be “interpretation” by the pathologist (for type 1 IHC results) for diagnostic

purposes or by the treating physician (for type 2 IHC results) for decisions directly relevant to patient treatment. With this approach, the readout is properly identified as a part of the analytical phase, just like it is in any other laboratory test (eg, nephelometry, mass spectrometry, etc.).<sup>66,67</sup> Furthermore, this approach also applies to IA; IA is intuitively accepted as “readout,” not as “interpretation” because it is performed by machines executing automated algorithms rather than by a living person.

## CONCLUSIONS

Biomarker testing by IHC is critical to patient care in the era of precision medicine. The IHC test is purpose-driven and purpose-defined. In today’s world, an IHC test is only meaningful when it is linked with a specific purpose for which it has been developed.

Part 1 of this 4-part series highlights the following:

1. The purpose of an IHC test is intrinsic to the test itself; if it is discovered that the current use of the test is different from its original purpose, the new purpose requires consideration of whether or not a new fit-for-purpose development and validation process is needed.
2. The result of an IHC test is not the IHC slide itself with the signal pattern that emerges on the tissue after development of the reaction product; rather, it is the readout of that signal pattern. This readout is performed by a pathologist, sometimes with the assistance of computer aided IA. Readout of IHC slides does require intellectual activity, observation and judgment, but this process should be considered as separate from interpretation of laboratory results in a strict sense, which is about what the result “means.”
3. The readout (test result) is a part of the analytical phase of IHC testing and is distinct from interpretation of the meaning, significance and utility of the test result. Generally, interpretation is a postanalytic test element performed by a pathologist (for diagnosis) or a treating physician ( $\pm$  pathologist) (for treatment decisions).
4. Using a fit-for-purpose approach, IHC tests can be classified as “type 1-IHC” tests or “type 2-IHC” tests based on the end-user.

## REFERENCES

1. Taylor CR. Predictive biomarkers and companion diagnostics. The future of immunohistochemistry: “in situ proteomics,” or just a “stain”? *Appl Immunohistochem Mol Morphol*. 2014;22:555–561.
2. de Matos LL, Trufelli DC, de Matos GL, et al. Immunohistochemistry as an important tool in biomarkers detection and clinical practice. *Biomark Insights*. 2010;5:9–20.
3. Taylor CR. Immunohistochemistry for the age of molecular morphology. *Appl Immunohistochem Mol Morphol*. 2001;9:1–2.
4. Gu J, Taylor CR. Practicing pathology in the era of big data and personalized medicine. *Appl Immunohistochem Mol Morphol*. 2014;22:1–9.
5. Taylor CR, Shi SR, Chaiwun B, et al. Strategies for improving the immunohistochemical staining of various intranuclear prognostic markers in formalin-paraffin sections: androgen receptor, estrogen

- receptor, progesterone receptor, p53 protein, proliferating cell nuclear antigen, and Ki-67 antigen revealed by antigen retrieval techniques. *Hum Pathol*. 1994;25:263–270.
6. Boreczuk AC, Allen TC. PD-L1 and lung cancer: the era of precision medicine? *Arch Pathol Lab Med*. 2016;140:351–354.
  7. Lee JW, Figeys D, Vasilescu J. Biomarker assay translation from discovery to clinical studies in cancer drug development: quantification of emerging protein biomarkers. *Adv Cancer Res*. 2007;96:269–298.
  8. Lee JW, Weiner RS, Sailstad JM, et al. Method validation and measurement of biomarkers in nonclinical and clinical samples in drug development: a conference report. *Pharm Res*. 2005;22:499–511.
  9. Pennello GA. Analytical and clinical evaluation of biomarkers assays: When are biomarkers ready for prime time? *Clin Trials*. 2013;10:666–676.
  10. Lee JW, Devanarayan V, Barrett YC, et al. Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm Res*. 2006;23:312–328.
  11. Rimm DL. What brown cannot do for you. *Nat Biotechnol*. 2006;24:914–916.
  12. Lohse J, Petersen KH, Woller NC, et al. Improved catalyzed reporter deposition, iCARD. *Bioconjug Chem*. 2014;25:1036–1042.
  13. van der Loss CM. Multiple immunoenzyme staining: methods and visualizations for the observation with spectral imaging. *J Histochem Cytochem*. 2008;56:313–328.
  14. Rudbeck L. Adding quality to your qualitative IHC| MLO. Available at: [www.mlo-online.com/adding-quality-qualitative-ihc](http://www.mlo-online.com/adding-quality-qualitative-ihc). Accessed July 24, 2016.
  15. United Nations Office on Drugs and Crime, Laboratory and Scientific Section. *Guidance for the Validation of Analytical Methodology and Calibration of Equipment Used for Testing of Illicit Drugs in Seized Materials and Biological Specimens: A Commitment to Quality and Continuous Improvement*. New York, NY: United Nations; 2009.
  16. Anon. Biomarkers in risk assessment: validity and validation (EHC 222, 2001). Available at: [www.inchem.org/documents/ehc/ehc/ehc222.htm](http://www.inchem.org/documents/ehc/ehc/ehc222.htm). Accessed July 24, 2016.
  17. Strimbu K, Tavel JA. What are biomarkers? *Curr Opin HIV AIDS*. 2010;5:463–466.
  18. Anon. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69:89–95.
  19. Jørgensen JT. Companion diagnostic assays for PD-1/PD-L1 checkpoint inhibitors in NSCLC. *Expert Rev Mol Diagn*. 2016;16:131–133.
  20. Ríos A, Barceló D, Buydens L, et al. Quality assurance of qualitative analysis in the framework of the European project “MEQUALAN”. *Accreditation Qual Assur*. 2003;8:68–77.
  21. Rocchetti TT, Silbert S, Gostnell A, et al. Validation of a multiplex real-time PCR assay for detection of *Mycobacterium* spp., *Mycobacterium tuberculosis* complex, and *Mycobacterium avium* complex directly from clinical samples by use of the BD max open system. *J Clin Microbiol*. 2016;54:1644–1647.
  22. van Krieken JHJM, Langerak AW, Macintyre EA, et al. Improved reliability of lymphoma diagnostics via PCR-based clonality testing: report of the BIOMED-2 concerted action BHM4-CT98-3936. *Leukemia*. 2007;21:201–206.
  23. Marchiò C, Dowsett M, Reis-Filho JS. Revisiting the technical validation of tumour biomarker assays: how to open a Pandora's box. *BMC Med*. 2011;9:1.
  24. Khleif SN, Doroshow JH, Hait WN, et al. AACR-FDA-NCI cancer biomarkers collaborative consensus report: advancing the use of biomarkers in cancer drug development. *Clin Cancer Res*. 2010;16:3299–3318.
  25. 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. *Arch Pathol Lab Med*. 2010;134:907–922.
  26. Taylor CR. Focus on biospecimens: the issue is the tissue. Observations on the “NCI-NIST fitness-for-purpose quality assessment and standards development workshop,” October 2010. *Appl Immunohistochem Mol Morphol*. 2011;19:95–98.
  27. O'Hurley G, Sjöstedt E, Rahman A, et al. Garbage in, garbage out: a critical evaluation of strategies used for validation of immunohistochemical biomarkers. *Mol Oncol*. 2014;8:783–798.
  28. Hanson CA, Ross CW, Schnitzer B. Anti-CD34 immunoperoxidase staining in paraffin sections of acute leukemia: comparison with flow cytometric immunophenotyping. *Hum Pathol*. 1992;23:26–32.
  29. Abenoza P, Lillemo T. CD34 and factor XIIIa in the differential diagnosis of dermatofibroma and dermatofibrosarcoma protuberans. *Am J Dermatopathol*. 1993;15:429–434.
  30. Abdalla S, Haboubi N, Kumar S. Quantification of microvessel density in human tumours. *Int J Oncol*. 1996;9:923–926.
  31. Mohammed ZMA, Orange C, McMillan DC, et al. Comparison of visual and automated assessment of microvessel density and their impact on outcome in primary operable invasive ductal breast cancer. *Hum Pathol*. 2013;44:1688–1695.
  32. Nielsen O, Nielsen S, Røge R, et al. Assessment Run 42 2014: Paired box gene 8 protein (PAX8). Available at: [http://nordiqc.org/downloads/assessments/42\\_64.pdf](http://nordiqc.org/downloads/assessments/42_64.pdf). Accessed July 24, 2016.
  33. Rodig SJ, Mino-Kenudson M, Dacic S, et al. Unique clinicopathologic features characterize ALK-rearranged lung adenocarcinoma in the western population. *Clin Cancer Res*. 2009;15:5216–5223.
  34. Gascoyne RD. ALK-positive diffuse large B-cell lymphoma is associated with Clathrin-ALK rearrangements: report of 6 cases. *Blood*. 2003;102:2568–2573.
  35. Cutz J-C, Craddock KJ, Torlakovic E, et al. Canadian anaplastic lymphoma kinase study: a model for multicenter standardization and optimization of ALK testing in lung cancer. *J Thorac Oncol*. 2014;9:1255–1263.
  36. Hapgood G, Savage KJ. The biology and management of systemic anaplastic large cell lymphoma. *Blood*. 2015;126:17–25.
  37. Wagner JA, Williams SA, Webster CJ. Biomarkers and surrogate end points for fit-for-purpose development and regulatory evaluation of new drugs. *Clin Pharmacol Ther*. 2007;81:104–107.
  38. Woodcock J. The prospects for “personalized medicine” in drug development and drug therapy. *Clin Pharmacol Ther*. 2007;81:164–169.
  39. Woodcock J, Woosley R. The FDA critical path initiative and its influence on new drug development. *Annu Rev Med*. 2008;59:1–12.
  40. Hall JA, Salgado R, Lively T, et al. A risk-management approach for effective integration of biomarkers in clinical trials: perspectives of an NCI, NCRI, and EORTC working group. *Lancet Oncol*. 2014;15:e184–e193.
  41. Taylor CR. FDA issues final rule for classification and reclassification of immunochemistry reagents and kits. *Am J Clin Pathol*. 1999;111:443–444.
  42. Food and Drug Administration/Center for Devices and Radiological Health. Guidance documents (medical devices and radiation-emitting products)—guidance for submission of immunohistochemistry applications to the FDA; final guidance for industry, 1998. Available at: [www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm094002.htm](http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm094002.htm). Accessed July 24, 2016.
  43. Anon. Study group 1 of the global harmonization task force. Principles of in vitro diagnostic (IVD) medical devices classification. 2007. Available at: [www.imdrf.org/docs/ghtf/archived/sg1/technical-docs/ghtf-sg1-n045r12-in-vitro-diagnostic-classification-070209.doc](http://www.imdrf.org/docs/ghtf/archived/sg1/technical-docs/ghtf-sg1-n045r12-in-vitro-diagnostic-classification-070209.doc). Accessed July 24, 2016.
  44. Canadian Association of Pathologists-Association canadienne des pathologistes National Standards Committee, Torlakovic EE, Riddell R, 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.
  45. Teutsch SM, Bradley LA, Palomaki GE, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) initiative: methods of the EGAPP working group. *Genet Med*. 2009;11:3–14.
  46. Anon. Health Canada. Medical devices bureau: draft guidance for the risk based classification system of in vitro diagnostic devices—1998 Health Canada consultation notice. Available at: [www.hc-sc.gc.ca/dhp-mps/md-im/applic-demande/guide-ld/ivd-rsk\\_idiv-rsq-eng.php](http://www.hc-sc.gc.ca/dhp-mps/md-im/applic-demande/guide-ld/ivd-rsk_idiv-rsq-eng.php). Accessed July 24, 2016.

47. Anon. Guide on medical devices (MD/IVD) CE marking (mark) & European (EEA/EU/EC) authorized representative service. Available at: [www.ce-marking.com/medical-devices.html](http://www.ce-marking.com/medical-devices.html). Accessed July 24, 2016.
48. Adams EJ, Green JA, Clark AH, et al. Comparison of different scoring systems for immunohistochemical staining. *J Clin Pathol*. 1999;52:75–77.
49. van Diest PJ, van Dam P, Henzen-Logmans SC, et al. A scoring system for immunohistochemical staining: consensus report of the task force for basic research of the EORTC-GCCG. European Organization for Research and Treatment of Cancer-Gynaecological Cancer Cooperative Group. *J Clin Pathol*. 1997;50:801–804.
50. Misch D, Blum T, Boch C, et al. Value of thyroid transcription factor (TTF)-1 for diagnosis and prognosis of patients with locally advanced or metastatic small cell lung cancer. *Diagn Pathol*. 2015;10:21.
51. Maniar KP, Wang Y, Visvanathan K, et al. Evaluation of microinvasion and lymph node involvement in ovarian serous borderline/atypical proliferative serous tumors: a morphologic and immunohistochemical analysis of 37 cases. *Am J Surg Pathol*. 2014;38:743–755.
52. Waterman TA, Hagen JA, Peters JH, et al. The prognostic importance of immunohistochemically detected node metastases in resected esophageal adenocarcinoma. *Ann Thorac Surg*. 2004;78:1161–1169.
53. Arigami T, Natsugoe S, Uenosono Y, et al. Lymphatic invasion using D2-40 monoclonal antibody and its relationship to lymph node micrometastasis in pN0 gastric cancer. *Br J Cancer*. 2005;93:688–693.
54. Storr SJ, Safuan S, Mitra A, et al. Objective assessment of blood and lymphatic vessel invasion and association with macrophage infiltration in cutaneous melanoma. *Mod Pathol*. 2012;25:493–504.
55. Laurinavicius A, Plancoulaine B, Laurinaviciene A, et al. A methodology to ensure and improve accuracy of Ki67 labelling index estimation by automated digital image analysis in breast cancer tissue. *Breast Cancer Res*. 2014;16:R35.
56. Choudhury KR, Yagle KJ, Swanson PE, et al. A robust automated measure of average antibody staining in immunohistochemistry images. *J Histochem Cytochem*. 2010;58:95–107.
57. Brüggmann A, Eld M, Lelkaitis G, et al. Digital image analysis of membrane connectivity is a robust measure of HER2 immunostains. *Breast Cancer Res Treat*. 2012;132:41–49.
58. Røge R, Riber-Hansen R, Nielsen S, et al. Proliferation assessment in breast carcinomas using digital image analysis based on virtual Ki67/cytokeratin double staining. *Breast Cancer Res Treat*. 2016;158:11–19.
59. Lange H. Digital pathology: a regulatory overview. *Lab Med*. 2011;42:587–591.
60. Bui M, Lloyd M, Allam-Nandyala P, et al. Using image analysis as a tool for assessment of prognostic and predictive biomarkers for breast cancer: How reliable is it? *J Pathol Inform*. 2010;1:29.
61. Taylor CR. The total test approach to standardization of immunohistochemistry. *Arch Pathol Lab Med*. 2000;124:945–951.
62. Park S-Y, Kim B-H, Kim J-H, et al. Panels of immunohistochemical markers help determine primary sites of metastatic adenocarcinoma. *Arch Pathol Lab Med*. 2007;131:1561–1567.
63. 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.
64. Watkins JCM, Nucci MR, Ritterhouse LL, et al. Unusual mismatch repair immunohistochemical patterns in endometrial carcinoma. *J Surg Pathol*. 2016;40:909–916.
65. Hickner J, Thompson PJ, Wilkinson T, et al. Primary care physicians' challenges in ordering clinical laboratory tests and interpreting results. *J Am Board Fam Med*. 2014;27:268–274.
66. Pasch A, Farese S, Gräber S, et al. Nanoparticle-based test measures overall propensity for calcification in serum. *J Am Soc Nephrol*. 2012;23:1744–1752.
67. Jungmann JH, MacAleese L, Buijs R, et al. Fast, High resolution mass spectrometry imaging using a medipix pixelated detector. *J Am Soc Mass Spectrom*. 2010;21:2023–2030.