

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/2066/173084>

Please be advised that this information was generated on 2019-06-01 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, 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 3: Technical Validation of Immunohistochemistry (IHC) Assays in Clinical IHC Laboratories

Emina E. Torlakovic, MD, PhD,* †‡ Carol C. Cheung, MD, PhD, JD,* § Corrado D'Arrigo, MB, ChB, PhD, FRCPath, ¶ ## Manfred Dietel, MD, PhD,** Glenn D. Francis, MBBS, FRCPA, MBA, FFS (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, ¶ ¶ ¶ Mogens Vyberg, MD, §§§ || || || Xiaoge Zhou, MD, #### ***** and Clive R. Taylor, MD, † † † † †

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

Abstract: Validation of immunohistochemistry (IHC) assays is a subject that is of great importance to clinical practice as well as basic research and clinical trials. When applied to clinical practice and focused on patient safety, validation of IHC assays creates objective evidence that IHC assays used for patient care are “fit-for-purpose.” Validation of IHC assays needs to be properly informed by and modeled to assess the purpose of the IHC assay, which will further determine what sphere of validation is required, as well as the scope, type, and tier of technical validation. These concepts will be defined in this review, part 3 of the 4-part series “Evolution of Quality Assurance for Clinical Immunohistochemistry in the Era of Precision Medicine.”

Key Words: biomarkers, quality assurance, quality control, technical validation, revalidation, immunohistochemistry

(*Appl Immunohistochem Mol Morphol* 2017;25:151–159)

In the last decade, the development of precision medicine and the high throughput discovery methods that support it have led to increasing use of selective biomarkers for diagnosis, prognosis, and prediction of response to targeted therapy.^{1–17} This has also led to increasingly stringent criteria for establishing and monitoring of test performance characteristics in biomarker testing, and has improved processes for validating methods that are used to detect and measure these biomarkers.¹⁸ The American Association for Cancer Research (AACR), Food and Drug Administration (FDA), and National Cancer Institute (NCI) formed the AACR-FDA-NCI Cancer Biomarkers Collaborative to accelerate the translation of novel cancer therapeutics into the clinic.¹⁸ The AACR-FDA-NCI consensus recommendations were designed to advance the use of biomarkers in cancer drug development, the harmonization of biomarker validation

Received for publication September 2, 2016; accepted October 3, 2016.

From the *Department of Laboratory Medicine and Pathobiology, University of Toronto; Departments of †Laboratory Hematology; §Pathology, 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, and Calgary Laboratory Services, Calgary, AB, Canada; ‡Canadian Immunohistochemistry Quality Control (CIQC), Vancouver, BC, Canada/Canadian Association of Pathologists National Standards Committee for High Complexity Testing/Immunohistochemistry; ||Poundbury Cancer Institute; ¶Dorset County Hospital NHS Foundation Trust, Dorchester, UK; #Cancer Diagnostic Quality Assurance Services (CADQASc), Dorchester; ***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, and Department of Clinical Medicine, Aalborg University, Aalborg University Hospital; ||||Nordic Immunohistochemistry Quality Control (NordiQC), Aalborg, Denmark; ####Department of Pathology, Beijing Friendship Hospital, Capital Medical University, Beijing; *****Chinese Committee for Pathologists—Immunohistochemistry Quality Control, China; and ††††Keck School of Medicine, University of Southern California, Los Angeles, CA.

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@utoronto.ca).

Copyright © 2017 Wolters Kluwer Health, Inc. All rights reserved.

and the development of reference standards and best practices for analytical validation.¹⁸

Immunohistochemistry (IHC) is an in situ immunoassay that detects targeted biomarkers in tissue samples. The role of IHC in clinical practice has become increasingly diverse and the advent of companion diagnostic testing to support precision medicine has only accelerated that trend.^{12,19–22} Similarly, its role in research has become increasingly important in the support of basic research and discovery and its translation to clinical practice through pre-clinical and clinical trials. IHC has many characteristics that make this assay both widely used and highly appreciated:

- its in situ nature enables detection of protein expression/distribution in different tissues, cell populations and in cell compartments;
- it can be highly sensitive and specific;
- it can be multiplexed;
- it is inexpensive, fast, and can be automated.

The most important challenges of the method are:

- it does not generate any results without human readout;
- its results are mostly descriptive;
- preanalytical conditions, essential to the outcomes of the testing, can be difficult to control or standardize in the clinical setting, where neither ischemic time nor conditions of fixation can necessarily be the same at all times.

This paper builds upon recently issued guidelines for validation of antibodies and clinical IHC assays, supports recommendations for the research use of biomarkers in clinical trials of anticancer drugs and offers a novel approach to our understanding of validation and verification for clinical applications of IHC. Although this paper describes concepts that are focused on IHC methodology, the principles are also immediately applicable to other in situ, cell-based assays.

TERMINOLOGY AND DEFINITIONS

Proper terminology is essential for defining validation in the context of IHC testing and describing how validation needs to be planned and executed. The most important definitions and clarifications relevant to IHC validation are:

- (1) Definition of “validation” for IHC testing.
- (2) Definition of “verification” for IHC testing.
- (3) Characteristics of IHC validation.
- (4) Critical components of the IHC test.
- (5) Who initiates and who performs the validation?
- (6) The role of external quality assurance (EQA) in validation.
- (7) The role of comparison with non-IHC methodologies for IHC validation.
- (8) Planning of Technical Validation of IHC testing

Definition of Validation of IHC Testing

Validation is defined as “confirmation, through the provision of objective evidence, that requirements for a specific intended use or application have been fulfilled” (ISO 9000).²³ “Intended use” or “intended application” are synonymous with “purpose.”

Provision of “objective evidence” is fulfilled by collecting evidence that the relevant test performance characteristics (TPCs) have been identified and assessed. Validation provides a “high degree of assurance” that these needs will be met (adapted from the Guideline on General Principles of Process Validation).^{24–26}

- Purpose: the first step in validation is establishing the purpose of the IHC test in question. The purpose is critically informed by the needs of the user. Consideration must also be given to whether or not the proposed IHC test is the appropriate one for this purpose. The parameters used in constructing the validation of an IHC test will determine whether that test is appropriate for its purpose (fit-for-purpose). Certain IHC tests are declared fit-for-purpose by regulatory agencies (The US FDA, among many others—see table 3 in part 1 of the Evolution series) after considering various types of evidence in the approval process.^{27–30} For laboratory developed tests (LDTs), evidence of whether a given IHC test is fit-for-purpose may be found in published literature in the form of diagnostic validation of specific marker(s) or clinical validation of a biomarker in clinical trials,^{31,32} as both reasonably require the prerequisite development and technical validation of an assay using the antibody intended for that use.
- Objective evidence: validation of testing requires planning for how to create objective evidence that a test will reproducibly meet the needs of the user. In clinical IHC testing, similar to any other laboratory test, this evidence is obtained from evaluation and documentation of TPCs.^{7,33,34} For TPCs relevant to IHC testing, see part 2 of the Evolution series.
- Level of assurance: the level of assurance that the needs of the user will be met is generally defined as “high degree of assurance” and relates to test accuracy. In practical terms, it refers to EQA accuracy and/or non-IHC methodology (NIM) accuracy (see part 2 of the Evolution series for definitions) and can be statistically deduced from the performance of the IHC assay against the known comparator. It is on this basis that the CAP Quality Center recommended the use of at least 20 cases with known positive or negative results in a technical validation set for a given type 1-IHC assay (assuming specific validation guidelines do not already exist), arguing that a sufficiently “high degree of assurance” (based on 95% confidence intervals) are achieved with agreements of $\geq 90\%$ between actual and expected values³⁵ The level of assurance (or accuracy) for a particular IHC test, has in some instances already been specifically addressed by professional organizations. Examples of the latter include the agreement of $\geq 95\%$ with alternative validated method or same validated method for HER2.²⁶

Definition of Verification of IHC Testing

Verification is defined as “confirmation, through the provision of objective evidence, that specified requirements have been fulfilled” (ISO 9000).²³

Verification requires confirmation that pre-determined specifications are consistently met, meaning that the test, an analyte or other IHC test component performs as expected from a technical perspective or that the laboratory is performing an already established test correctly. “Established” refers to those tests that were previously either diagnostically and/or clinically validated for a specific purpose. In other words, verification confirms that the reagents supplied and the protocols used yield the appropriate analytical sensitivity, specificity, and reproducibility. In the “fit-for-purpose” framework, verification of an IHC assay is tier 1 technical validation (see below tiers of validation).

All IHC kits cleared or approved by the FDA and other regulatory agencies need to be verified before being put into clinical use, because while the supplier can provide proof that the kit is designed and performing well if properly used, the supplier cannot ensure that the laboratory performing the test will actually properly use the kit. Therefore, the manufacturer is responsible for validation (manufacturer provides evidence that the test they are selling is appropriate test for the condition to be demonstrated), but not for verification. Verification or demonstration that a validated test is achieving the desired analytical sensitivity, specificity, and reproducibility is the responsibility of the laboratory performing the test. If the manufacturer/supplier provides exact instructions and tools for verification, they should be followed.

Characteristics of Validation

Validation can only be properly designed for any given IHC assay after all characteristics of validation are considered in the planning stage. There are 4 characteristics of validation: sphere, scope, type, and tier.

- Spheres of validation: there are 3 spheres of validation: technical, diagnostic, and clinical. Clinical IHC laboratories typically perform only technical validation. Therefore, the rest of the article will refer to technical validation unless otherwise stated. In principal, “technical validation” is required for all IHC tests (whether they are applied to investigational or clinical questions) and it is the IHC laboratory that performs technical validation of these protocols. Although clinical IHC laboratories generally do not perform “diagnostic validation” or “clinical validation” of IHC tests, evidence for diagnostic validation or clinical validation is required for all tests that will be performed in the clinical IHC laboratory. Such evidence is usually available in peer-reviewed published literature. Neither CLIA’88 nor CAP requires clinical or diagnostic validation to be done by the IHC laboratory before clinical implementation of an IHC assay. Therefore, reliance on the evidence of diagnostic and clinical validations in published peer-reviewed literature is appropriate.
- Scope of technical validation: scope of technical validation (or validation of the analytical phase of IHC testing) includes initial validation and (when appropriate) revalidation. Initial validation is required

when a test is first introduced as well as when a new purpose is articulated for an already validated test. Revalidation is required when: (1) there is a change in a critical test component (primary revalidation; for the definition of “critical test components” see below); (2) when there is a change in a test performance characteristic that occurred without any change to critical test components (secondary revalidation); and (3) when there is a change in noncritical test components only (tertiary revalidation). This is summarized in Table 1. One of the most common reasons for primary revalidation is a change in the primary antibody. For example, for a CD3 IHC test that has been validated for detection of T-cell differentiation with a monoclonal antibody that is no longer commercially available, the introduction of a new monoclonal or polyclonal anti-CD3 antibody requires primary revalidation because the test is now functionally an entirely new assay, irrespective of its commercial name. When the supplier has changed, but the primary antibody is of the same clone, only tertiary revalidation is required. So that these changes are transparent to the end user, both the laboratory test method and the pathology reports that detail the readout and interpretation of that test, different clones and different polyclonal antibodies should be referred to by both their clone name (or generic name for a the polyclonal antibody), and by their molecular target. Although using gene names may appear to be universally acceptable, antibodies are developed for the proteins and, on occasion, gene name may be less informative (eg, antibodies for phosphorylated proteins or mutated gene products).^{36–38} Epitopes detected by different antibodies may have different performance characteristics related to preanalytical parameters. Some CD3 epitopes may be resistant to prolonged cold ischemic times, while others may degrade very quickly after devitalization. Some epitopes significantly change during fixation and embedding and may not be recovered to the same degree as others with a given antigen retrieval protocol.^{39–42} In principle, IHC assays based on different antibodies against a common target (such as CD3) are different assays that must be validated separately before their introduction to clinical testing.³⁵ Recognition that there is a need for primary revalidation is particularly important for type 2-IHC tests [eg, estrogen receptor (ER) for breast cancer, anaplastic lymphoma kinase for lung cancer, etc.]. If a new clone for ER for breast cancer or anaplastic lymphoma kinase for lung cancer is introduced for clinical use in the laboratory that is already using some other ER or anaplastic lymphoma kinase clone, this is another example that represents a change in a critical test component and thus requires primary revalidation. In contrast, secondary revalidation is required when there is observed change in any of the relevant test performance characteristics; these may occur either due to reproducible change in results with on-slide controls or suboptimal results in proficiency testing (PT) (the latter only if there is evidence of previous success in PT by using the same protocol).

TABLE 1. Scope of Technical Validation

Scope	Trigger	Tiers
Initial	New test or new purpose	Tier 2 or 3
Revalidation (no new purpose)		
Primary	Change in critical test components (\pm change in TPCs)	Tier 2 or 3
Secondary	Change in TPCs, but no change in critical test components	Tier 1 for type 1-IHC Tier 2 for type 2-IHC
Tertiary	Change in noncritical test components only (eg, changes in primary Ab dilution)	Tier 1

Ab indicates antibody; IHC, immunohistochemistry; TPCs, test performance characteristics.

- Type of technical validation: when the term “validation” is used, it is essential to be specific about what is being validated as this defines what type of technical validation is being performed. In IHC, validation may refer to: generic validation of primary antibody, validation of the same clone from a different supplier, lot-to-lot primary antibody validation, validation of buffers, diluents, visualization system, and other reagents, validation of equipment, validation of preanalytical conditions (fixative, fixation time, decalcifying reagents, decalcification time, etc.), validation of controls, technical validation of IHC protocols of biomarkers, and validation of IHC readout.^{5,43} This series addresses specifically 2 types of validation: (i) technical validation of IHC protocols of biomarkers (Fig. 1), and (ii) validation of the IHC readout (Fig. 2).
- Tiers of technical validation: tiers of technical validation relate to TPCs included in validation. Technical validation spans from tier 1 to tier 3. Tier 3 validation (full validation) consists of determining analytical sensitivity and specificity, analytical reproducibility, demonstration of reportable range, determining desired (predetermined) high level of accuracy within reportable range, determining protocol preanalytical robustness, and determining readout accuracy and precision where applicable. Figure 1 summarizes all 3 tiers of technical validation of the IHC protocol. Validation of the entire analytical phase is completed by inclusion of validation of the readout (Fig. 2). The tissue tools required for different tiers of validation are described in part 4 of the Evolution series. Ideally, all IHC tests performed by an accredited clinical laboratory for patient care should have evidence of tier 3 validation; however, this is not always possible. Evidence gathered in tier 2 validation is essential for clinical applications of all IHC tests and applies to initial validation of IHC LDTs.

Critical Components of the IHC Test

IHC is a complex process, the results of which are influenced by multiple parameters in the preanalytical,

analytical, and postanalytical test phases. Potential critical components in all of these steps must be identified and defined based on published evidence along with cumulative clinical and EQA experience. Defining the critical components of clinical laboratory tests is important because if changes are made to any critical component of an already validated IHC test, then primary revalidation of the IHC test needs to be performed. This is why “in-house” changes to an FDA (or other regulatory agency) cleared or approved IHC test automatically transform it into a LDT and, as such, the test no longer carries the imprimatur of the regulatory agency.⁴² Similarly, introducing changes to an already validated LDT requires consideration regarding the need for possible revalidation.

- (a) Preanalytical phase: While the preanalytical phase is difficult to control, it is reasonably well-defined and its importance is generally agreed upon.⁴⁴⁻⁴⁸ There are numerous components in the preanalytical phase.⁴⁸ Of these, a subset can be identified for which changes would automatically trigger the need for primary revalidation of the IHC protocol. The critical components of the preanalytical phase include: (1) fixative type, (2) decalcifying reagent type, and (3) instrumentation for tissue processing. Any change in other preanalytical components (ie, noncritical components), while it may not automatically trigger the need for primary revalidation, may still ultimately require secondary or tertiary revalidation if the noncritical component change leads to unacceptable alterations in test performance characteristics.
- (b) Analytical phase: Because of intricacies related to its in situ, cell-based, descriptive nature, as well as the dual role of the pathologist (readout and interpretation), the components of the analytical phase of the IHC test are complex and are not immediately self-evident. The analytical phase for formalin-fixed/paraffin-embedded tissues consists of four major components: (1) antigen retrieval method, (2) primary antibody, (3) detection system, and (4) readout (Fig. 3). Although there are other components of the analytical phase, it is any of these four critical components that will necessitate primary revalidation of the IHC protocol if they are changed. When critical components (1), (2), and (3) are combined, they are synonymous with “IHC protocol.”⁴⁹ The addition of component 4 (readout) completes the analytical phase (see part 1 from Evolution series) (Fig. 2). Although ultimately the IHC slides will be sent to pathologists for readout and interpretation, the design of the readout validation is under the purview of the IHC laboratory medical director who ensures that readout methods are defined and align with the purpose of the test. Any change in other analytical components (ie, noncritical components), while it may not automatically trigger the need for primary revalidation, may still ultimately require secondary or tertiary revalidation if the noncritical component change leads to unacceptable alterations in monitored test performance characteristics.

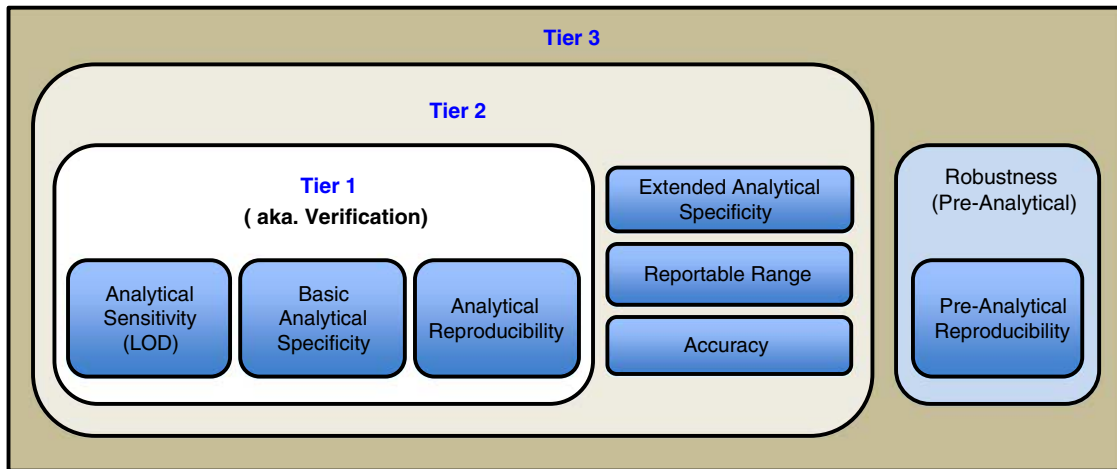


FIGURE 1. Three tiers of validation of the immunohistochemistry protocol. Each tier of technical validation is based on the type of evidence it provides for the test validation (by including different test performance characteristics). LOD indicates limit of detection.

(c) Postanalytical phase: The postanalytical phase includes interpretation of the readout by the end user for a particular intended use. This is where the “use” of the test results should align with the “purpose” for which the test was validated. If the use of the test results is not linked to the purpose for which the test was originally validated, then initial validation needs to be performed to determine whether the current IHC protocol can fulfill the new purpose (see part 1 of the Evolution series), or whether a new IHC protocol needs to be developed and validated for the new purpose. The use may be altered by either pathologists (for type 1-IHC tests) or by the treating physician (for type 2-IHC tests). In such situations, by changing the use of the test, the end user is effectively relying on an

unvalidated IHC testing the same way that a treating physician is using an “unvalidated” clinical management plan by administering a medical therapy “off-label.” It is important to recognize that such a decision with respect to change of “use” of an IHC test may be unknown to the laboratory and thus, the responsibility to initiate validation lies with the end user (pathologist or oncologist).

Who Initiates and Who Performs the Validation?

The responsibility for the validation of various components of the IHC test lies with the party that controls those components. Responsibility for IHC test

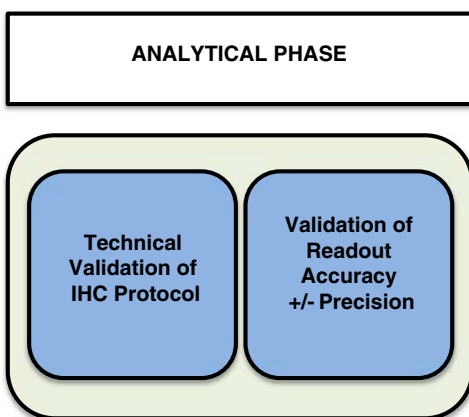


FIGURE 2. Validation of IHC analytical phase includes validation of IHC protocol (Fig. 2) as well as validation of the readout. Readout validation consists of demonstrating that readout accuracy and readout precision are acceptable for clinical practice. If image analysis is used as an aid for the readout, then it also needs validation to demonstrate a desirable level of accuracy and precision. IHC indicates immunohistochemistry. full color online

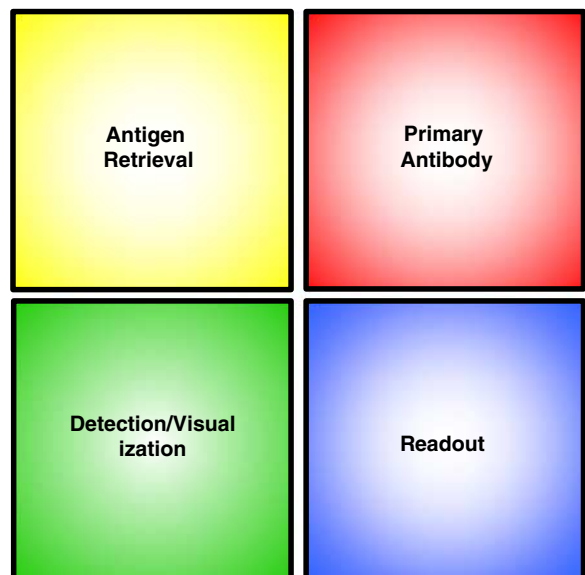


FIGURE 3. “Four squares” of the immunohistochemistry analytical phase, which starts with antigen retrieval (including “none” as an option) and is completed by the readout. full color online

validation can be divided into: (i) responsibility for initiation of IHC test validation, and (ii) responsibility for design and conducting the IHC test validation.

Responsibility for initiation of validation of various IHC test components is delegated to their most appropriate “owner.” The clinical IHC laboratory and its medical director are responsible for triggering the initial validation of the analytical phase whenever a new IHC test is being introduced or for triggering revalidation of the IHC protocol whenever changes in the critical components of the IHC protocol or its intended use are made, or unacceptable change in monitored test performance characteristics occurs as discussed earlier. It must be emphasized that the responsibility for the validation of the readout lies solely with the medical director of the laboratory. In addition, if a critical component of an IHC test is being made in areas that are not controlled by the IHC laboratory, the responsible party needs to consider the implications of the changes for that IHC test. This also includes the responsibility of the pathologist (other than medical director of IHC laboratory) to trigger the validation of the IHC test if he/she changes the readout of the IHC test or the use of the IHC test in such way that it does not align with the originally validated purpose of the test. Otherwise, the responsibility for validation of various IHC test components is as follows:

- purchased reagents need to be validated by the suppliers who make claims about the reagents relevant to its “fit-for-purpose” principle;
- purchased or leased equipment needs to be validated by the suppliers who make claims about the equipment; although clinical laboratories often perform some sort of verification of the new equipment (and may be helped in this process by the supplier), full validation of the equipment is done by the supplier;
- “home-made” reagents and equipment need to be validated by the party that put the “home-made” reagent/equipment into use;
- initial or changed preanalytical components need to be validated in collaboration with surgical pathology (ie, grossing room), histology and the IHC laboratory;
- pathologist interpretation of type 1-IHC tests for diagnostic use (diagnostic validation of type 1-IHC) is validated in the published literature and if not already published, internal diagnostic validation is required by the pathologist who is using the IHC test or by the IHC medical director;
- clinical validation for predictive and prognostic markers is generally published in the literature; this is typically based on the results of clinical trials and other clinical studies (or rarely it is established internally when appropriate follow-up or other clinical data are available), although not all published “clinical validations” meet the rigorous criteria initially proposed by McGuire;⁵⁰
- FDA-approved (or other regulatory agency cleared or approved) IHC kits are already clinically validated for a specific purpose; however, verification (tier 1 validation) of these kits remains the responsibility of the IHC laboratory utilizing the kits.

It is clear from the above that the validation of IHC testing requires a team approach and that the IHC laboratory can only be responsible for validation of the components of the test over which it has direct control, though it must be prepared to respond appropriately to internal evidence that intended use and readout parameters have been or are being modified in local practice.

The Role of EQA in IHC Validation

EQA has an important role in IHC validation because it may provide important information about various IHC TPCs. It can provide an evidentiary basis for IHC test analytical sensitivity and specificity (ie, whether the test was properly calibrated), IHC test accuracy (ie, agreement of results with designated standard), accordance (ie, the probability that 2 identical test materials analyzed by the same laboratory under standard repeatability conditions will both yield the same result; for example, such that both are found to be positive or both are found to be negative), concordance (ie, the probability that 2 identical test materials sent to different laboratories will both yield the same result).^{51–54} It can be performed by participation in PT or by interlaboratory testing as appropriate.^{55,56} EQA programs often use a small number of tissue samples that are immunohistochemistry critical assay performance controls (iCAPCs) or their equivalent for PT. The results of such PT challenges usually provide good evidence of analytical sensitivity and specificity.^{15,57–59} Indeed, well-controlled PT exercises can generate a high level of evidence as a basis for defining TPCs. EQA programs or interlaboratory testing that include TMA samples with the sufficient number and proper selection of tissue samples can provide evidence of test accuracy (EQA accuracy).^{60,61} When EQA samples include tissues that have been subjected to various known preanalytical conditions, they provide evidence of preanalytical IHC protocol robustness. Some EQA programs also provide clinical validation of IHC testing. This is rare, but it has been accomplished by the German EQA program (QuIP—Quality in Pathology) for HER2 in gastric cancer as the PT sample design had been developed by the ToGA Study – A Study of Herceptin (Trastuzumab) in Combination With Chemotherapy Compared With Chemotherapy Alone in Patients With HER2-Positive Advanced Gastric Cancer.⁶²

The Role of Comparison With Non-IHC Methods for IHC Validation

For some IHC tests, comparison with results of non-IHC methodologies can provide evidence of IHC protocol accuracy (NIM accuracy; see part 2 of the Evolution series). The alternative methodology needs to be assessed as to whether it has been already validated for the same intended clinical application and whether there is evidence that it can be considered as the “designated true value.” The alternative methodology should not be used for IHC validation if it has not been properly validated for the same purpose. The comparison can be performed intralaboratory, interlaboratory, or by participation in PT. If the laboratory can internally evaluate

concordance with non-IHC methodology results, external validation may not be necessary for tier 2 validation. This approach is typically used in tier 2 validation for type 2-IHC tests (eg, HER2 IHC validation by fluorescence in situ hybridization).²⁶

Planning of Technical Validation for IHC Testing

Validation of IHC testing requires a formal validation plan, which needs to be completed before operational implementation. At the planning stage, all aspects of technical validation need to be considered to determine what scope, type, and tier of validation need to be executed. In other words, whether it is an initial validation or revalidation that is required, what is being validated, and which TPCs will be evaluated for the validation need to be formulated. Furthermore, a “fit-for-purpose” approach requires that any test be validated for specific intended use with an identified end user, as well as identified needs of the end user. This planning and preparation also involves the creation of proper quality

tools (as a component of the laboratory’s quality management plan) that will be needed not only for test development, but also for on-going quality monitoring (see part 4 of the Evolution series). Table 2 illustrates steps of the planning stage for initial technical validation. A similar approach may also be feasible for revalidation.

CONCLUSIONS

- “Validation” provides objective evidence at a high level of assurance that a test is fit-for-purpose.
- “Validation” should always be used with a qualifier describing the sphere, scope, type, and tier of validation.
- “Verification” is synonymous with tier 1 technical validation.
- “Revalidation” should always be used with a qualifier describing whether it is primary, secondary, or tertiary and what trigger prompted the reevaluation.

TABLE 2. Planning of Initial Technical Validation of IHC Testing

Steps That Need Documentation	Illustrative Examples	
	Type 1-IHC (eg, Melan-A)	Type 2-IHC (eg, ER)
1. Identify/designate the name of the test	Melan-A	Estrogen Receptor (ER)
2. Identify the purpose of the IHC test*	Detection of melanocytic differentiation	Detection of ER expression in breast cancer
3. Identify end user	Pathologist	Oncologist
4. Identify needs of the end user	The need for a highly sensitive and specific marker of melanocytic differentiation at the diagnostic sensitivity and specificity demonstrated in published literature	The need for quantitative marker that will predict response to tamoxifen at the rate demonstrated in previously published clinical trials
5. Identify if the marker has been diagnostically and/or clinically validated†	Yes (properly diagnostically validated in published literature)	Yes (properly clinically validated in published literature)
6. Identify if there is relevant peer-reviewed published evidence on technical validation	No (details of technical validation not available)	Yes (details of technical validation available)
7. Select publication(s) that will guide protocol development‡	List relevant PubMed or other peer-reviewed publications	List relevant PubMed or other peer-reviewed publications
8. Identify if previously validated primary Ab is available commercially‡	Yes (MART-1) or other as relevant	Yes (SP-1) or other as relevant
9. Identify if tissues required for test development and controls are available	Search revealed that adrenal gland, kidney, strongly positive malignant melanoma, weakly positive melanoma, and ovarian granulosa cell tumor are useful (eg, www.nordiqc.org)	Search in published literature revealed that cervix, tonsil, and 4 breast carcinomas with graded predetermined expression range from 0% to 100% are useful as well as 50 positive and 50 negative cases of breast carcinoma (eg, www.nordiqc.org, ASCO/CAP guidelines)
10. Identify what is required for the new test readout‡	Qualitative assessment by pathologist	Qualitative and quantitative assessment by pathologists; image analysis may be applied
11. Identify appropriate EQA program or external laboratory and/or non-IHC method for tier 2 validation	EQA program that provides Melan-A challenges or any laboratory with demonstrated success in EQA testing for Melan-A (laboratories with expertise in melanoma diagnostics are preferred); could be used for EQA accuracy	EQA program that provides PT samples for ER challenges or designated reference laboratory for ER; if PT samples contain sufficient number of breast carcinoma cases, they may be used for EQA accuracy
12. Develop tools for test development and on-going quality monitoring	See part 4 of Evolution Series Develop iCAPCs, Index and Accuracy TMA for type 1-IHC	See part 4 of Evolution Series Develop iCAPCs, Index and Accuracy TMA for type 2-IHC

*It is assumed that the purpose is clinically justifiable.
 †If the marker has not been diagnostically and/or clinically validated, the laboratory cannot use the test or the laboratory needs to perform full diagnostic and/or clinical validation as appropriate; IHC medical director or pathologist(s) that request new test development should provide relevant publications.
 ‡If primary antibody is not commercially available or tissues are not available, it may not be possible to develop proper test.
 Ab indicates antibody; ASCO/CAP, American Society of Clinical Oncology/College of American Pathologists; ER, estrogen receptor; EQA, external quality assurance; IHC, immunohistochemistry; iCAPCs, immunohistochemistry critical assay performance controls; PT, proficiency testing.

- The IHC laboratory is responsible for technical validation of all IHC tests that are performed in the IHC laboratory and must be prepared to produce objective evidence of test performance (ie, by evaluating relevant test performance characteristics) at the relevant levels of assurance (as suggested by professional organizations or accrediting bodies).
- Planning of technical validation starts with the consideration of what scope, type, and tier of validation need to be executed.
- Use the same primary antibodies that have previously been found to be “fit-for-purpose” and introduce/develop and technically validate IHC protocols that can reasonably reproduce results appropriate to previously published or disseminated diagnostic or clinical validation studies. If an IHC test has not been properly diagnostically or clinically validated in published literature and it will be introduced for diagnostic or other clinical purposes, the IHC laboratory needs to perform a diagnostic and/or clinical validation as appropriate. For most clinical laboratories this is not possible, so these tests should not be performed.
- The tests that are cleared or approved by the FDA or other regulatory bodies were already validated for a specific purpose and therefore only need verification by the end-user laboratory (tier 1 validation) if a different method of verification is not specifically instructed by the manufacturer or by professional practice guidelines.
- Comparison with IHC results from PT and/or inter-laboratory collaboration as well as non-IHC method results can be used to determine IHC protocol accuracy and robustness, which are important for tier 2 validation.

REFERENCES

- 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.
- Dunstan RW, Wharton KA, Quigley C, et al. The use of immunohistochemistry for biomarker assessment—can it compete with other technologies? *Toxicol Pathol*. 2011;39:988–1002.
- 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.
- 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 (unabridged version). *Arch Pathol Lab Med*. 2010;134:e48–e72.
- Howat WJ, Lewis A, Jones P, et al. Antibody validation of immunohistochemistry for biomarker discovery: recommendations of a consortium of academic and pharmaceutical based histopathology researchers. *Methods*. 2014;70:34–38.
- 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.
- 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.
- 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.
- Mengel M, Hebel K, Kreipe H, et al. Standardized on-slide control for quality assurance in the immunohistochemical assessment of therapeutic target molecules in breast cancer. *Breast J*. 2005;11:34–40.
- 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.
- Pennello GA. Analytical and clinical evaluation of biomarkers assays: when are biomarkers ready for prime time? *Clin Trials*. 2013;10:666–676.
- 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.
- Taylor CR, Levenson RM. Quantification of immunohistochemistry—issues concerning methods, utility and semiquantitative assessment II. *Histopathology*. 2006;49:411–424.
- 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.
- 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.
- Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69:89–95.
- World Health Organization and International Programme on Chemical Safety. *Biomarkers in Risk Assessment: Validity and Validation*. Geneva: WHO; 2001. Available at: www.inchem.org/documents/ehc/ehc/ehc222.htm. Accessed July 24, 2016.
- 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.
- Conde E, Hernandez S, Prieto M, et al. Profile of Ventana ALK (D5F3) companion diagnostic assay for non-small-cell lung carcinomas. *Expert Rev Mol Diagn*. 2016;16:707–713.
- Jørgensen JT. Companion diagnostic assays for PD-1/PD-L1 checkpoint inhibitors in NSCLC. *Expert Rev Mol Diagn*. 2016;16:131–133.
- Sholl LM, Aisner DL, Allen TC, et al. Programmed death ligand 1 immunohistochemistry—a new challenge for pathologists: a perspective from Members of the Pulmonary Pathology Society. *Arch Pathol Lab Med*. 2016;140:341–344.
- Taylor CR. Immunohistochemistry in surgical pathology: principles and practice. *Methods Mol Biol*. 2014;1180:81–109.
- International Organization for Standardization. ISO 9000:2015(en), Quality management systems—fundamentals and vocabulary. 2015. Available at: www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v1:en. Accessed August 27, 2016.
- Baseman H. Commentary on the US Food and Drug Administration’s 2011 “Guidance for Industry, Process Validation General Principles and Practices”, in: Rathor AS and Sofer G. Process Validation in Manufacturing of Biopharmaceuticals. CRC Press 2012; 11–50. Available at: www.crcnetbase.com/doi/pdf/10.1201/b12013-3. Accessed August 27, 2016.
- 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.
- Wolff AC, Hammond MEH, Schwartz JN, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol*. 2006;25:118–145.
- Health Canada. Medical devices bureau: draft guidance for the risk based classification system of in vitro diagnostic devices—1998 Health Canada Consultation Notice. 2016. Available at: www.hc-sc.gc.ca/dhp-mps/md-im/applie-demande/guide-ld/ivd-rsk_idiv-rsq-eng.php. Accessed July 24, 2016.

28. European Commission. In vitro diagnostic medical devices. 1998. Available at: https://ec.europa.eu/growth/single-market/european-standards/harmonised-standards/iv-diagnostic-medical-devices_en. Accessed August 28, 2016.
29. European Commission. Revisions of medical device directives—growth. 2012. Available at: http://ec.europa.eu/growth/sectors/medical-devices/regulatory-framework/revision_en. Accessed August 28, 2016.
30. US Food and Drug Administration. 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. Accessed July 24, 2016.
31. Miettinen M, Wang Z-F, Lasota J. DOG1 antibody in the differential diagnosis of gastrointestinal stromal tumors: a study of 1840 cases. *Am J Surg Pathol*. 2009;33:1401–1408.
32. Ma D, Wang Z, Yang L, et al. Responses to crizotinib in patients with ALK-positive lung adenocarcinoma who tested immunohistochemistry (IHC)-positive and fluorescence in situ hybridization (FISH)-negative. *Oncotarget*. 2016;7:64410–64420.
33. International Council for Harmonization. Validation of analytical procedures: text and methodology. Q2 R1. 2005. Available at: <http://somatek.com/content/uploads/2014/06/sk140605h.pdf>. Accessed August 8, 2016.
34. Burd EM. Validation of laboratory-developed molecular assays for infectious diseases. *Clin Microbiol Rev*. 2010;23:550–576.
35. 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.
36. Kurosawa N, Wakata Y, Inobe T, et al. Novel method for the high-throughput production of phosphorylation site-specific monoclonal antibodies. *Sci Rep*. 2016;6:25174.
37. Gao W, Yang J, Liu W, et al. Site-specific phosphorylation and microtubule dynamics control Pyrin inflammasome activation. *Proc Natl Acad Sci*. 2016;113:E4857–E4866.
38. Anwar MAF, Murad F, Dawson E, et al. Immunohistochemistry as a reliable method for detection of BRAF-V600E mutation in melanoma: a systematic review and meta-analysis of current published literature. *J Surg Res*. 2016;203:407–415.
39. Shi S-R, Shi Y, Taylor CR. Antigen retrieval immunohistochemistry: review and future prospects in research and diagnosis over two decades. *J Histochem Cytochem*. 2011;59:13–32.
40. Shi SR, Cote RJ, Taylor CR. Antigen retrieval immunohistochemistry: past, present, and future. *J Histochem Cytochem*. 1997;45:327–343.
41. Shi SR, Cote RJ, Taylor CR. Antigen retrieval techniques: current perspectives. *J Histochem Cytochem*. 2001;49:931–937.
42. Taylor CR. Standardization in immunohistochemistry: the role of antigen retrieval in molecular morphology. *Biotech Histochem*. 2006;81:3–12.
43. Bordeaux J, Welsh A, Agarwal S, et al. Antibody validation. *BioTechniques*. 2010;48:197–209.
44. Portier BP, Wang Z, Downs-Kelly E, et al. Delay to formalin fixation “cold ischemia time”: effect on ERBB2 detection by in-situ hybridization and immunohistochemistry. *Mod Pathol*. 2013;26:1–9.
45. Sato M, Kojima M, Nagatsuma AK, et al. Optimal fixation for total preanalytic phase evaluation in pathology laboratories. A comprehensive study including immunohistochemistry, DNA, and mRNA assays: optimal fixation. *Pathol Int*. 2014;64:209–216.
46. Plebani M, Sciacovelli L, Aita A, et al. Harmonization of pre-analytical quality indicators. *Biochem Med*. 2014;24:105–113.
47. Plebani M, Sciacovelli L, Aita A, et al. Quality indicators to detect pre-analytical errors in laboratory testing. *Clin Chim Acta*. 2014;432:44–48.
48. Taylor CR. The total test approach to standardization of immunohistochemistry. *Arch Pathol Lab Med*. 2000;124:945–951.
49. Taylor CR, Shi S-R. Techniques of immunohistochemistry: principles, pitfalls, and standardization. In: Dabbs DJ, ed. *Diagnostic Immunohistochemistry*. Philadelphia, PA: Elsevier Health Sciences; 2013:1–38.
50. McGuire WL. Breast cancer prognostic factors: evaluation guidelines. *J Natl Cancer Inst*. 1991;83:154–155.
51. Lord RS, Bralley JA. *Laboratory Evaluations for Integrative and Functional Medicine*. Duluth, GA: Metamatrix Institute; 2008.
52. Langton SD, Chevenement R, Nagelkerke N, et al. Analysing collaborative trials for qualitative microbiological methods: concordance and concordance. *Int J Food Microbiol*. 2002;79:175–181.
53. van der Voet H, van Raamsdonk LWD. Estimation of concordance and concordance in inter-laboratory trials of analytical methods with qualitative results. *Int J Food Microbiol*. 2004;95:231–234.
54. Copete M, Garratt J, Gilks B, et al. Inappropriate calibration and optimisation of pan-keratin (pan-CK) and low molecular weight keratin (LMWCK) immunohistochemistry tests: Canadian Immunohistochemistry Quality Control (CIQC) experience. *J Clin Pathol*. 2011;64:220–225.
55. Bashkansky E, Gadrach T, Kuselman I. Interlaboratory comparison of test results of an ordinal or nominal binary property: analysis of variation. *Accred Qual Assur*. 2012;17:239–243.
56. International Standards Organization/International Electrotechnical Commission. ISO/IEC 17025:2005(en). General requirements for the competence of testing and calibration laboratories. 2005. Available at: www.iso.org/obp/ui/#iso:std:iso-iec:17025:ed-2:v1:en. Accessed August 29, 2016.
57. NordiQC—Immunohistochemical Quality Control. PAX8. 2014. Available at: <http://nordiqc.org/epitope.php?id=64>. Accessed July 24, 2016.
58. Williams J, Parry S. The general pathology module Run 113. Immunocytochemistry. April 3–13, 2016.
59. Vyberg M, Nielsen S. Proficiency testing in immunohistochemistry—experiences from Nordic Immunohistochemical Quality Control (NordiQC). *Virchows Arch*. 2016;468:19–29.
60. Pérez T, Makretsov N, Garatt J, et al. Modeling Canadian quality control test program for steroid hormone receptors in breast cancer: diagnostic accuracy study. *Appl Immunohistochem Mol Morphol*. 2016;24:679–687.
61. Makretsov N, Gilks CB, Alaghehbandan R, et al. Development of an evidence-based approach to external quality assurance for breast cancer hormone receptor immunohistochemistry: comparison of reference values. *Arch Pathol Lab Med*. 2011;135:874–881.
62. Bang Y-J, Van Cutsem E, Feyereislova A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet*. 2010;376:687–697.