

Best Laboratory Practice for Examination of Prognostic and Predictive Markers by Immunohistochemistry



Prof. Emina Emilia Torlakovic, MD, PhD, FCAP
UHN/Toronto General Hospital
Canada

Objectives

- Background
- Terminology
- Classification of IHC tests
- Current recommendations
- Future of Class II IHC

Regulation of Clinical Laboratory Tests in the US

- In 1988, USA Congress passed the Clinical Laboratory Improvement Amendment (CLIA) to establish quality standards for all aspects of clinical laboratory testing.
- FDA has primary responsibility to classify tests based on 1 of 3 CLIA regulatory categories: waived for simple tests, and moderate or high complexity for other tests based on their potential risk to public health.
- CDC serves an advisory role and provides scientific support.

Regulation of Clinical Laboratory Tests in the US

- Code of Federal Regulations (CFR) is a compilation of the general and permanent rules published in the Federal Register by various departments and agencies of the US government.
- Title 21 is reserved for FDA rules.
- In 1997, FDA published 3 rules that were intended to ensure the quality of and control the risks associated with reagents used in laboratory developed tests (LDT).
- Analyte Specific Reagents (ASR) were introduced, which protected LDTs from overregulation.

LDT, ASR, “Home Brew Tests”

- All IHC tests except FDA approved kits.
- All molecular tests except FDA approved kits.
- All cytogenetic tests and FISH.
- “CLIA regulated laboratories qualified to perform high complexity testing have demonstrated expertise and ability to use ASRs in test procedures and analyses”.

Requirements for Test Validation

- The FDA does not specify requirements for test validation, but it provides guidance for commercial manufacturers that intend to submit validation data for FDA approval or clearance.
- When FDA-approved kits are used, laboratory only need to confirm its performance characteristics (verify the test claims).

LDT Validation

- The laboratory must establish test performance specifications:
 - Accuracy
 - Precision
 - Reportable range
 - Reference range
- The laboratory must develop and plan procedures for calibration and control of the test system.
- The laboratory must establish analytical sensitivity and specificity.

Principles of Test Validation

- ISO 9000 – “Confirmation by using objective evidence, that requirements for a specific intended use or application have been fulfilled.”
- Validation – we are doing the *correct* test.
- Verification – we are doing the test *correctly*.
- Validation requires *identification of the needs of the user*.

Analytic Performance Characteristics

- **Accuracy**: our result – reference value (or conventional true value) = error
- **Trueness**: systematic error/bias
- **Precision** (for quantitative tests): measure of random error (SD)
- **Reproducibility** (precision)
- **Repeatability**: reproducibility within-run
- **Reference range**: range of test values for designated population

Analytic Performance Characteristics

- **Analytic sensitivity**: positive agreement as compared to reference method
- **Analytic specificity**: negative agreement as compared to reference method
- **Clinical sensitivity**: proportion of subjects with a disorder with positive test result
- **Clinical specificity**: proportion of subjects without disorder with negative test result
- **Limit of detection**: the lowest amount of analyte (Ag) that is statistically distinguishable from background or negative control

Cochrane Collaboration

- 2003 **Standards for the Reporting of Diagnostic Accuracy**
- **Diagnostic accuracy**: agreement between test results and reference standard.
- **Reference standard**: the best available method for establishing the presence or absence of the condition of interests.
- **Reference standard**: single method, combination of methods, imaging, pathology, clinical follow-up, etc.



CBC News in Depth:



Misdiagnosed - Anatomy of Newfoundland's Cancer-Testing Scandal

- Of the 1,013 breast cancer patients retested (1997-2005), 383 — more than a third — were found to be false negative. That meant 383 patients were denied a fighting chance against cancer. More than 100 of those wrongly tested patients are now dead.
- What's more, in another cruel twist, it would later come out that not all of those affected were even notified that a mistake had been made.

What Went Wrong in Newfoundland? How to Fix it?

- Media responds with various takes on the subject.
 - “Breast Cancer Testing Scandal Shines Spotlight on Black Box of Clinical Laboratory Testing” JNCI News
- What really went wrong? Who was responsible and why?

Commission of Inquiry on Hormone Receptor Testing

The Honourable Justice Margaret A. Cameron, Commissioner

- The Commission expressed conclusions and recommendations regarding responsibility of various persons or organizations, and delivered its final report and recommendations to the Minister of Health and Community Services on before **February 28, 2009.**
- QA
- QA
- QA

<http://www.cihrt.nl.ca/about.html>

AJCP / SPECIAL ARTICLE

Canadian Association of Pathologists—Association canadienne des pathologistes National Standards Committee/Immunohistochemistry

Best Practice Recommendations for Standardization of Immunohistochemistry Tests*

Emina Emilia Torlakovic, MD, PhD,¹ Robert Riddell, MD, FRCPath, FRCPC,² Dipsankar Banerjee, MBChB, FRCPC, PhD,³ Hala El-Zimaity, MD, MS, FRCPC,⁴ Dragana Pilavdzic, MD, FRCPC,⁵ Peter Dawe, MS,⁶ Anthony Magliocco, MD, FRCPC,⁷ Penny Barnes, MD, FRCPC,⁸ Richard Berendt, MD, FRCPC,⁹ Donald Cook, MD, FRCPC,¹⁰ Blake Gilks, MD, FRCPC,¹¹ Gaynor Williams, MD, PhD,¹² Bernardo Perez-Ordóñez, MD, FRCPC,¹³ Bret Wehli, MD, FRCPC,¹⁴ Paul E. Swanson, MD,¹⁵ Christopher N. Otis, MD,¹⁶ Soren Nielsen, HT, CT,¹⁷ and Mogens Vyberg, MD,¹⁷ and Jagdish Butani, MBBS, MS, FRCPC¹³

Key Words: Clinical immunohistochemistry; Standards; Canadian Association of Pathologists

DOI: 10.1093/AJCP/21/06/414

Guidelines:

Table of Contents

- Use of Standard Terminology in Clinical Immunohistochemistry
- Principles/Best Practices for Quality Assurance of Clinical IHC Testing
- Class II Immunohistochemistry Tests Principles/Best Practices
- Proficiency testing: Monitoring the quality of laboratory performance
- Education and training standards for laboratory personnel
- References

Current Role of IHC in Pathology

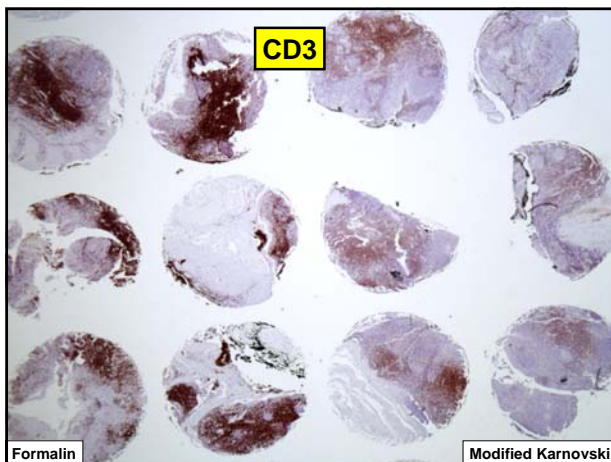
- Limitation of morphology alone
- Limitation of molecular or cytogenetic studies in relation to morphological and immunophenotypic evidence
- Limitation of flow cytometry – in situ demonstration may be more informative
 - Greatly dependent on the type of the lesion
 - Focal lesions may be missed by flow
 - Rare lesional cells may be missed if the diagnosis is not already established

Qualitative IHC

- IHC has been in use in diagnostic pathology for about 4 decades
- It is only in the last 10 years that technological advances enabled us to claim that:
- **IHC results are highly reproducible**
- **IHC can be finely tuned/calibrated, and**
- **IHC is amenable to standardization**

Literature Review: Methods' Description is Insufficient

	Not Described (%)	Significance for IHC Results
Fixation	27	Essential
Fixation time	75	
Decalcification	52	Essential
Decalcification time	72	
Positive controls	87	Very important
Negative controls	80	Very important
Pretreatment methods	48	Essential
Pretreatment buffer	61	Essential
Pretreatment time	68	Essential
Detection system	17	Essential



Standardization vs. Optimization

- **Pre-Analytical variables of IHC tests** – Any and all steps in tissue processing, including intraoperative tissue handling/treatment (prolonged ischemia, delayed fixation, etc.), type and length of fixation, decalcification, and elements of tissue handling. The pre-analytical component is concluded at microtomy and the placement of the tissue section on pre-treated glass slides.
- **Analytical variables of IHC tests** – The analytical variables phase begins with the handling of the cut slides in a clinical IHC laboratory. It is completed with the coverslipping of the stained slides.
 - Antibodies, controls, automation, reagents
- **Post-Analytical variables of IHC tests** – Interpretation and reporting of the results.

Challenges in Standardization of Immunohistochemistry

- **Standardization of protocols is meaningless without control standardization.**
- **“Standardization” is greatly misused term in this context.**
 - Unfortunately, diagnostic IHC is not fully amenable to true standardization.
 - As long as tissue processing cannot be fully standardized, diagnostic IHC can be only optimized.
 - However, this appears to be highly desirable at the current state of art diagnostic IHC.

CAP-ACP IHC Tests Classification (Clinical Use)

- **Class I** – Used by pathologists
- **Class II** – Used by clinicians for the purpose of patient treatment and follow up

Class I

- The results are incorporated into the diagnostic interpretation by the pathologists.
- Cytokeratin, S-100, vimentin, CD45, CDX2, TTF1, etc.

Classification Relates to Risks to Health

- Misdiagnosis and initiation of inappropriate therapies or withholding of appropriate therapies
- The degree of risk depends on whether the product is used as an adjunct to conventional histopathological diagnostic techniques or provides information that is used independently.
- **Based on FDA approach to classification of IHC devices and reagents (relevant to industry)**

Class II IHC tests

- Despite the need for finely tuned calibration and quantitative nature of the tests, they are usually reported simply as positive or negative.
- The simplicity of the report masks the true biological and technical complexity of the testing.

Class II IHC

- Stand alone diagnostic
- Predictive or prognostic
- Generally, results are interpreted irrespective of histology or other IHC results
- E.g. hormone receptors and HER2 in breast cancer

Class II IHC Tests

- **Prognostic IHC tests** – The results of these tests independently forecast clinical outcome. They may be either qualitative or quantitative. HER2/neu if used as prognostic marker.
- **Predictive IHC tests** – The results of these tests independently predict response to a particular therapy. They may either be qualitative or quantitative (e.g., ER/PR, HER2/neu in breast carcinoma, CD117 in gastrointestinal stromal tumor).

Interpretation and Test Classification

- Class I IHC tests, which have critical significance for interpretation of overall assessment:
ALK-1, cyclin D1, CD30, TdT, TTF-1, CDX-2, HMB-45, ...
- New IHC test Class may be necessary to raise awareness and prevent wrong diagnoses.

Challenges

- There are about 200 IHC tests that are currently in clinical use.
 1. Most are not included in proficiency testing
 2. Standardized controls are not available
 3. National and international agreement on what standardized controls should be does not exist even for Class II markers
 4. **Validation** of most IHC assays is not clearly defined:
 - For any calculations of sensitivity, specificity, and agreement power analysis should be considered so that calculations are not misleading (for some tests this may mean close to 100 samples to test).

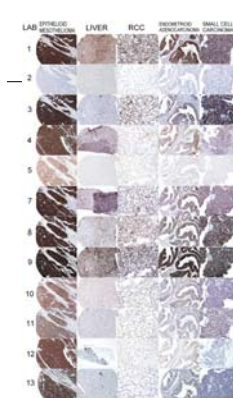
Lack of Quality Control Samples and Standardization

- Lack of definitions and/or agreement what samples should be used for either positive or negative controls
- Lack of actual source of QC samples
- Lack of funds for appropriate generation of QC samples
- Lack of standardized calibrators
- Lack of **knowledge dissemination in QC** including both laboratory physicians, technologists, managers, and users (oncologists, other...)

External Quality Assurance and Proficiency Testing (PT) in IHC

- Does not exist for many tests!
- Various programs that are providing PT do not clearly define their targets:
 - **What are gold standards?**
 - **What are reference values?**
 - **Are assessments quantitative for quantitative IHC tests?**
 - **Is participation sufficient to validate protocols?**
 - **Are they testing analytical or clinical sensitivity and specificity or both?**
 - **Are they addressing at all sensitivity and specificity?**

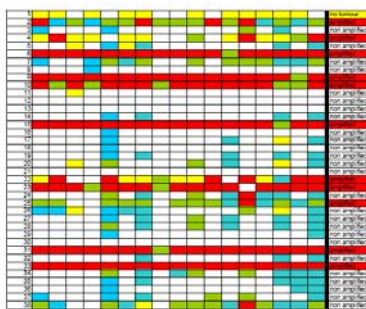
NordiQC and CIQC Experience



33% optimal
33% good/suboptimal
33% poor

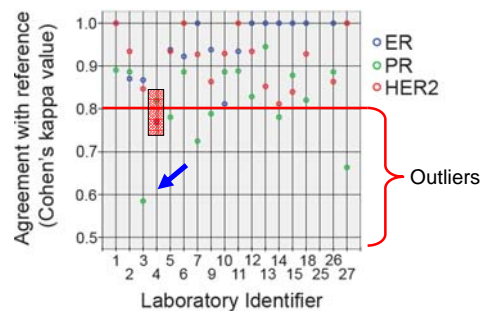
CIQC Run2: HER2

Average sensitivity: 91%, specificity: 98%



Kendall's coefficient of concordance = 0.96

How to Define Discordant Results?



High Expectations of Accuracy: To Treat or Not Treat?

- Sensitivity and specificity of most tests is not defined
- When possible to calculate sensitivity and specificity, standards are not set or not universally agreed upon
- There is no tracking of clinical impact of reported IHC tests
- There is no tracking of clinical impact of proficiency testing (PT) results of various programs that provide PT

Future

- Patient safety is clearly identified as the no. 1 priority in the design and regulation of laboratory testing by all agencies and organizations
- Standardization needs to address all parameters of the testing (pre-analytical, analytical, and post-analytical)
- QA measures need to be tailored to the test type (Class I vs. Class II) as well as to make biological and statistical sense.

Future: Era of -OMICS

- In situ demonstration of protein expression
- -Omics studies at the moment are relatively expensive and discovery-focused
- IHC required to confirm data obtained by other methods including standard NB, WB, and SB
- Omics studies narrow our focus from large scale to small scale (most important genes)
- Many “most important genes” are being detected by immunohistochemistry

Best Laboratory Practice

- Follow the rules! (If they exist...)
- Class II IHC tests should be performed and reported according to national and international guidelines
- You can go beyond laboratory accreditation requirements; poor regulatory practice does not excuse for not following the rules of best practice (if available)
- Nobody will accuse you of taking too much care of patient safety

Recommended Approach to Test Validation

- Planning –
 - What type of IHC test is it? What is intended use of the test? Class I or Class II?
 - Quantitative or qualitative?
 - Are there national or international guidelines?
 - What controls need to be used? Are they available?
 - Set acceptance criteria.
- Pilot testing, troubleshooting, and optimization

Generate Validation Data

- Evaluate positive and negative agreement with established standard or method.
- Evaluate reproducibility repeating the test on the same samples.
- Assess clinical performance (using patient samples with and without condition/Ag)

Clinical and Laboratory Standards Institute (CLSI) Recommendations: Evaluation Protocols

- Generally:
 - 10 to 20 operating days
 - 20 to 40 patient samples
 - 50 positive and 50 negative

- www.clsi.org/

Components of a Validation Report

- **Specimens tested**
- **Method comparison**
- **Analytic performance characteristics:** sensitivity, specificity, and (for quantitative tests) accuracy, precision, reproducibility, linearity, and analytic measurement range
- **Define:** reportable range, acceptable sample types, criteria for rejection, controls, standards, calibrators, step-by-step procedure, interpretation, and reporting guidelines.

What to do?

- Review your current QA systems for clinical IHC for its adequacy
- Proactively build appropriate internal and external QA measures to support development and clinical applications of new IHC tests
- Use appropriate positive and negative controls with inbuilt quantitative information
- Search for or contribute to agreement/consensus on IHC test classification within your discipline and with your clinical colleagues
- Identify key components according to published guidelines (and common sense!) that are not addressed by laboratory accreditation.
- Proactively improve how you report IHC test results for both Class I and Class II tests

Internal Controls: vWF (F8-ra)

