Introduction to Immunohistochemistry

Clive R. Taylor, MD, D.Phil

Immunohistochemistry (n.)
Microscopic localization of specific antigens in tissues by staining with antibodies labeled with fluorescent or pigmented material.

The American Heritage® Medical Dictionary
Chapter 1.1 | Introduction

Immunohistochemistry (IHC) is a method used to determine the expression of biomarkers in tissue. This educational guidebook will describe immunohistochemistry as it is used in the pathology laboratory as an aid in the differential diagnosis and classification of cancer, and for certain other diseases, including infections. The factors that influence the immunohistochemical staining result start in the surgery operating room and end at the interpretation of the stain by the pathologist, which ultimately leads to treatment decision by the oncologist.

For those new to the world of immunohistochemistry here is a brief outline of the steps needed to localize antigens in tissues using antibodies for cancer diagnosis:

Figure 1.1 From biopsy to reporting.
Pre-Analytical Steps
1. A **Biopsy** (surgically removed tissue specimen or needle biopsy) from the surgery room arrives in fixative at the pathology laboratory.
2. In the **Accessioning** room the sample details are entered into the laboratory information system (LIS). A barcoded label can ensure track and trace capabilities.
3. During **Grossing**, the specimen is visually examined for suspicious areas that require further examination. Samples from the specimen that require further microscopic testing are excised as tissue blocks and placed in barcoded cassettes.
4. **Tissue processing and embedding** are the steps where the tissue block is processed into a form and condition suitable for making ultrathin microscopic sections. Typically, the tissue is fixed in formalin then dehydrated before it is embedded in paraffin.
5. **Sectioning** is the fine art of cutting the paraffin-embedded tissue blocks into ultrathin (~4 µm) sections and placing them onto glass slides. A barcode on the slide can ensure traceability and may also contain protocol information for the requested test for that particular section.

Analytical Steps
6. **Staining** is the analytical part of the IHC process. It encompasses antigen retrieval, application of the primary antibody and visualization system, ending with counterstaining:
   a. Antigen retrieval is performed to recover the antigens that may have been altered by fixation;
   b. Endogenous enzymes are blocked (this step can also be performed after primary antibody incubation);
   c. A primary antibody is applied that specifically binds to the antigen of interest;
   d. The secondary antibody carries the label (enzyme); upon application it binds to the primary antibody;
   e. Chromogen is applied to visualize the antibody/antigen complex;
   f. Counterstaining is performed to visualize nuclei and overall tissue architecture;
   g. Sections are dehydrated, mounted and coverslipped.

Post-Analytical Steps
7. In the post-analytical process, the pathologist interprets the stains in context with positive and negative tissue controls, using bright field microscopy.
8. The results are reported to the oncologist for treatment decision.

Figure 1.2 Many factors may influence the IHC staining result. With just 3 choices at each of 14 steps there are 4.8 million different procedures!
Chapter 1.2 | History of immunohistochemistry

This IHC Educational Guidebook will describe the potentials and pitfalls in the immunohistochemical staining process from biopsy to interpretation, with special attention to the analytical processes and how to improve certainty in the staining result by employing standardization to the processes.

Before immunohistochemistry reached its now widespread use as an important method in routine cancer diagnosis, the technology had a long history of technological developments outlined in the table below.

Table 1.1 The major milestones in the history of immunohistochemistry.

<table>
<thead>
<tr>
<th>Year</th>
<th>Method</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1941</td>
<td>Fluorescence-labeled primary antibodies</td>
<td>Coons et al (1)</td>
</tr>
<tr>
<td>1967</td>
<td>Enzyme-labeled primary antibodies</td>
<td>Nakane &amp; Pierce (2)</td>
</tr>
<tr>
<td>1970</td>
<td>Secondary unlabeled antibodies (PAP)</td>
<td>Sternberger et al (3)</td>
</tr>
<tr>
<td>1974</td>
<td>Application to routine formalin paraffin sections</td>
<td>Taylor et al (5-7)</td>
</tr>
<tr>
<td>1975</td>
<td>Invention of monoclonal antibodies</td>
<td>Köhler &amp; Milstein (8)</td>
</tr>
<tr>
<td>1978</td>
<td>Double staining using unlabeled antibodies (APAAP)</td>
<td>Mason &amp; Sammons (9)</td>
</tr>
<tr>
<td>1979</td>
<td>Monoclonal antibodies to human antigens</td>
<td>McMichael et al (10)</td>
</tr>
<tr>
<td>1988</td>
<td>Capillary gap semi-automated staining</td>
<td>Brigati et al (11)</td>
</tr>
<tr>
<td>1991</td>
<td>Heat-induced antigen retrieval</td>
<td>Shi et al (12)</td>
</tr>
<tr>
<td>1993</td>
<td>Standardization efforts as ‘Total Tests’</td>
<td>Taylor (13)</td>
</tr>
<tr>
<td>1995</td>
<td>Dextran-polymer-based detection system</td>
<td>Dako</td>
</tr>
<tr>
<td>1998</td>
<td>Immunohistochemistry as companion diagnostics</td>
<td>Dako (HER2)</td>
</tr>
<tr>
<td>2008</td>
<td>Molecular HER2 CISH Tests in the IHC lab</td>
<td>Invitrogen</td>
</tr>
</tbody>
</table>

Professor Albert H. Coons and co-workers demonstrated in 1941 that it was possible to localize antigens in tissue slices using antibodies against Streptococcus pneumoniae labeled with fluorescein and visualized by ultraviolet light (fluorescence microscopy) (1). During the next 25 years, the Coons method was used with different modifications, including labeling with heavy metals, but it was not until the introduction of enzyme-labeled antibodies (2) that the method overcame many of the inherent issues with fluorescein and heavy metal labeling of antibodies. In the early 1970s, application of the ‘immunoperoxidase’ method to formalin paraffin embedded tissues by Taylor, Mason and colleagues in Oxford, was a critical step in extending use of the method into ‘routine’ diagnosis in anatomic pathology. The direct labeling method had the drawback that each individual primary antibody, or the secondary antibody, had to be labeled with enzyme. That problem was circumvented by the development of an unlabeled antibody enzyme method, the peroxidase anti-peroxidase (PAP) method, which had the further advantage of increased sensitivity, facilitating use in routine tissues. A related parallel development was the introduction of the alkaline phosphatase anti-alkaline phosphatase (APAAP) in 1978 (9). Even with the development of new and improved detection systems for visualization of antigens in tissue, IHC suffered from lack of reproducibility, due in part to poor quality antibody reagents, and in part to the inconsistent and adverse effects of fixation.

Increased demand led to better quality reagents from the commercial sector, with improved quality control of production methods. Polyclonal antibody preparations differ between serum samples in affinity and specificity, as the immune-response changes with time and immunization preparations, and as one animal is replaced by another as the source. Dr. Niels Harboe, founder of Dako, realized in the early 1970s the need for standardized antibody preparations for safe and reproducible diagnoses and began producing purified polyclonal antibodies that had the same strength (as measured by titer) from batch to batch.

Even with the purified and highly specific polyclonal antibodies there was a need for improved specificity of antibodies and a greater variety in terms of target proteins. The invention, in 1975, of hybridomas that could produce monoclonal antibodies (8) resulted in the production of the first monoclonal antibody that was highly specific for human thymocytes using hybridoma technology (10). Monoclonal antibodies paved the way for a rapid growth in the use of IHC in research and diagnosis of cancer.
One other consequence of the lack of reproducibility was the development of automated instruments (11). Automation was invented with the fundamental thought that a properly functioning and maintained instrument will consistently perform its pre-programmed instructions in the same way—slide after slide—which is the principal reason why an instrument potentially can give superior reproducibility, compared with manual methods. However, progress was slow until 1991, when Shi et al (12) introduced ‘antigen retrieval’ (or heat-induced epitope retrieval), thereby facilitating extension of IHC to a much broader range of applications in formalin paraffin sections, but at the cost of adding yet another variable to the process. This important publication on antigen retrieval thus gave new insights and impetus to efforts in standardization of IHC, leading to the introduction of the ‘Total Test’ concept (13) as a result of a series of meetings sponsored by the Biological Stain Commission and the FDA in the early 1990s.

The standardization efforts, coupled with attempts to use IHC in a semi-quantitative setting raised demands to a new level, exemplified by the introduction, in 1998, of the HercepTest™ (Dako), which was the first cancer companion diagnostic, in this instance designed for selection of breast cancer patients for treatment with the new drug Herceptin® (Genentech/Roche). Clinical trials had shown that patients whose tumors overexpressed HER2 would benefit the most from Herceptin® treatment. The HercepTest™ assay uses IHC on patient samples, in combination with control cell lines having known HER2 expression to determine if a breast cancer overexpresses HER2. Some 15 years later, this assay together with similar HER2 assays from other vendors, still serves as a rare example of a semi-quantitative IHC assay used in routine clinical pathology. The polymer-based visualization system, introduced shortly before HercepTest™, is the most widely used detection method in IHC today, with advantages of stability and high sensitivity.

The technical advances in IHC in the last decade have been incremental, with little impact on the basics of the method. Automation has become more advanced, including laboratory information system integration, with track and trace of samples, while whole slide digital imaging is slowly being integrated into the analysis of stain result. These advances can best be regarded as improvements in standardization; a process that started back

Figure 1.3 The development of detection systems used for IHC. Please see Chapter 6 for a full description of the many different detection methods.
in the early 1990s and was re-emphasized in the 2007 publications by Goldstein et al. (14) and Wolff et al. (15), but also by the work being done e.g. estrogen receptor assessment (16, 17). The critical importance of IHC standardization became evident with the revelation of disturbingly high numbers of false negative or false positive results in IHC determinations of ER (estrogen receptor) and PR (progesterone receptor) expression, and also HER2. In one example, a re-testing in 2007 of 1,023 breast cancer samples from Newfoundland revealed that approximately 1 out of 3 samples was scored falsely ER negative (17). As a consequence of the false negative ER test results, these women were not accorded the potential benefit of anti-hormonal therapy.

The latest development in cancer diagnosis is the inclusion of molecular tests (FISH/CISH) in anatomic pathology labs, driven by HER2 assessment requirements. Other technologies also are entering into the pathology lab and into routine diagnosis, and technologies such as array comparative genomic hybridization or next generation sequencing will likely be a fundamental part of cancer diagnosis in the future. One ongoing goal is to interface these newer methods of molecular analysis with existing and improved morphologic criteria, a field termed ‘Molecular Morphology’.

Immunohistochemistry is based on principles similar to the ELISA method, yet it is at best a semi-quantitative method for determination of the expression of biomarkers in tissue samples. However, IHC should not be regarded as simply another ‘special stain’, like a PAS stain or a silver stain. IHC is essentially an ELISA method applied to a tissue section. In this respect, when correctly performed, IHC has the potential to perform as a reproducible and quantitative tissue-based ELISA assay; much more than a simple stain. That the IHC method does not perform to this level, reflects deficiencies in the application of the method, specifically inconsistent sample preparation, lack of reference or calibration standards, and inadequate validation of reagents (18, 19). If ELISA can use a standard curve to convert the measured immunoreactivity into a quantitative amount of tested protein, then IHC – in theory – can also convert the IHC intensity observed in FFPE tissue sections into the amount of tested protein by an equivalent standard ruler. Comparative studies of IHC intensity on frozen tissue vs. FFPE tissue have shown identical intensity by using an optimized AR protocol (20, 21), and similar protein quality is evident when examined by mass spectrometry (22), leaving no theoretical reason for lack of true quantitative IHC assays. Nonetheless, today IHC assays are at best no more than semi-quantitative, for reasons that are more of a practical nature.

Figure 1.4 The number of IHC publications in the last 50 years. The data are from Pubmed using the search term “immunohistochemistry”.

**Chapter 1.3 Standardization in Clinical Immunochemistry vs. Immunohistochemistry**

For more than 30 years, clinical immunochemistry has employed blood or urine samples to determine the concentration of certain biomarkers, e.g. creatinine and cystatin C for evaluation of kidney function, and C-reactive protein as a marker of inflammation. Although clinical immunochemistry covers a multitude of assay types, most of these tests are based on the ELISA (enzyme-linked immunosorbent assay) method, a method that closely parallels IHC in principle. One major difference is that International Reference Materials and Calibrators are used in clinical immunochemistry (ELISA) to achieve quantitative results from these assays.
Chapter 1.4 | Growing Consensus for Standardization

From the beginning there has been concern relating to the reproducibility of immunohistochemical methods as applied to formalin-fixed, paraffin-embedded (FFPE) tissue sections. A consequence of not controlling all parameters (in fixation, processing and staining) is poor day to day reproducibility within a single laboratory, and poor reproducibility among different laboratories. In recent years these concerns have increased and lack of standardization, well shown in inter-laboratory quality assurance surveys performed by NordiQC and UK NEQAS, is now recognized as a major impediment to basic research, clinical trials, and direct patient care. Over the past three decades a number of conferences have been held to address this topic and to seek constructive resolutions. Among the most productive were a series of meetings sponsored by the Biological Stain Commission and the FDA in the early 1990s, that led to recommendations for manufacturers concerning the precise description and validation of IHC reagents (23), and also highlighted the necessity to pay attention to all aspects of the IHC test procedure. The latter recommendation, borrowed from the much more rigorous protocols applied to immunologic assays in clinical laboratories, became known as the ‘Total Test’ approach (Table 1) (23, 24). A decade later a meeting of the FDA and NIST (National Institute of Standards and technology) focused upon standardization of HER2 IHC assays, and the need for universal control materials (reference standards) (25).

Chapter 1.5 | Standardization Starts in the Surgery Room

While Table 1.2 only mentions a few of the major steps in a Total Test, the pre-analytical process alone contains at least 62 identifiable steps of which 27 have been examined in published research. Out of these 27 steps, 15 pre-analytical variables are capable of impacting the immunohistochemistry staining result including fixation delay, fixative type, time in fixative, reagents and conditions of dehydration, clearing, paraffin impregnation and conditions of slide drying and storage (26). Pre-analytical variables are described in detail in Chapter 2.

Table 1.2 The Total Test: An IHC stain should be managed in the same rigorous manner as a clinical laboratory analysis. Modified from Taylor (14, 24).

<table>
<thead>
<tr>
<th>Pre-analytic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test selection</td>
</tr>
<tr>
<td>Specimen type</td>
</tr>
<tr>
<td>Acquisition, pre-fixation/transport time</td>
</tr>
<tr>
<td>Fixation, type and total time</td>
</tr>
<tr>
<td>Processing, temperature</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antigen retrieval procedure</td>
</tr>
<tr>
<td>Selection of primary antibodies</td>
</tr>
<tr>
<td>Protocol; labeling reagents</td>
</tr>
<tr>
<td>Reagent validation</td>
</tr>
<tr>
<td>Control selection</td>
</tr>
<tr>
<td>Technician training/certification</td>
</tr>
<tr>
<td>Laboratory certification / QA programs</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Post-analytic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment of control performance</td>
</tr>
<tr>
<td>Description of results</td>
</tr>
<tr>
<td>Interpretation/reporting</td>
</tr>
<tr>
<td>Pathologist, experience and CME specific to IHC</td>
</tr>
</tbody>
</table>

In the analytical steps, antigen retrieval is the first challenge. Different antigens require different antigen retrieval for optimal staining results, and the different variations of the AR process adds another variable that must be controlled. Antigen retrieval is described in detail in Chapter 3.

Selecting the right antibody for the right marker is one of the key steps in the analytical process. Some monoclonal antibody clones are more specific than others against the same biomarker. In other cases a polyclonal antibody may be the best choice. Selection of the primary antibody is described in detail in Chapter 4.

Using a protocol that is optimized for the detection of the biomarker is vital. The optimal protocol must be able to identify the antigen of interest in cells and structures with both low and
high expression. Optimization of the staining protocol is described in detail in Chapter 5.

The final step of the analytical process is the visualization of the antigen/antibody reaction. Here the selection of the detection system must consider the complexity of the visualization and the required amplification needed to visualize the biomarker. The various detection systems are described in detail in Chapter 6.

Post-analytical standardization is essential for prognostic or predictive biomarkers, e.g. HER2 and ER/PR, adhering to specified stain interpretation guidelines to give the sample a scaled score (e.g. from 0-3+). However, most biomarkers are used for cell lineage and tissue identification, where expression levels are usually not as critical and interpretation is not linked to a semi-quantitative scoring system, but is reported as a binary ‘Yes’ or ‘No’ system (positive or negative) for the tested biomarker. Digital analysis of IHC stains is described in Chapter 7.

The consensus arising from the standardization efforts is that the reliability and reproducibility of IHC methods in routine surgical pathology have been greatly hindered by two key factors.

1. While reagents available for IHC have increased in quality, there has been an even greater increase in number of sources and variety of staining methods. This plentitude of reagents contributes to lack of standardization in significant ways, that in theory are manageable by good technique and use of proper controls, but in practice have led to requirements for such high standards of excellence in the technical process, that many laboratories cannot find sufficient, or sufficiently skilled, staff to comply.

2. The usual method of sample preparation for tissue remains as formalin fixation and paraffin embedment (FFPE). This venerable approach may be satisfactory for the preservation of morphologic detail, but does adversely affect the antigenicity of many target molecules in the tissue, to degrees that are unknown. The enormous variation in protocols (including fixation times) employed for FFPE among different laboratories, or within the same laboratory from specimen to specimen, compounds the problem and contributes to the current poor reproducibility.

While several decades have passed, these issues have not been satisfactorily addressed. Legions of investigators, and many manufacturers, have addressed different aspects of the problem, focusing upon better sample preparation (fixation), more effective methods of antigen retrieval, improved

**Table 1.3 Major steps affecting the immunohistochemistry staining result.**

<table>
<thead>
<tr>
<th>Step</th>
<th>Effect on IHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biopsy</td>
<td>Depending on the suspected cancer type, tissue samples can be obtained in different ways such as punch/core biopsy, excisional/incisional biopsy, etc. Tissue degradation begins at the time of sample removal.</td>
</tr>
<tr>
<td>Fixation</td>
<td>The sample should be fixed as soon as possible after surgery, ideally within less than an hour. The chemical fixation crosslink proteins in the sample thereby stopping the degradation process. Too short or too long fixation can affect the staining result.</td>
</tr>
<tr>
<td>Embedding</td>
<td>After fixation, the sample is embedded in paraffin for long-term storage and to enable sectioning for subsequent staining. Once embedded in paraffin, samples can be stored (almost) indefinitely.</td>
</tr>
<tr>
<td>Sectioning and Mounting</td>
<td>Formalin-fixed, paraffin-embedded tissues are sectioned into thin slices (4-5 μm) with a microtome. The sections are then mounted onto adhesive-coated glass slides.</td>
</tr>
<tr>
<td>Antigen Retrieval</td>
<td>Due to the fixation process, an antigen retrieval treatment is applied to unmask the epitopes, either by heat (heat-induced epitope retrieval; HIER) or enzymatic degradation (proteolytic-induced epitope retrieval; PIER). Incorrect antigen retrieval for the biomarker of interest will adversely affect the staining result.</td>
</tr>
<tr>
<td>Primary Antibody</td>
<td>An antibody with specificity for the biomarker of interest is applied. The specificity and sensitivity of the antibody affect the staining result.</td>
</tr>
<tr>
<td>Visualization</td>
<td>The antigen/antibody complex signal is amplified and visualized using a detection system. The strength of amplification of the reaction affects the staining result (intensity).</td>
</tr>
<tr>
<td>Interpretation</td>
<td>The staining pattern is assessed by a pathologist in context with other biomarkers, controls and other tests (e.g. H&amp;E, special stains. Inter- and intra-observer variability is common, especially for semi-quantitative assays. This variability highlights the importance of training and inter-calibration.</td>
</tr>
</tbody>
</table>
reagents, more sophisticated automated platforms, more sensitive detection methods, and the development of reference standards or controls (13, 23-25).

In order to improve the quality and reproducibility from sample to sample, and lab to lab, the accreditation process for many pathology laboratories now includes participation in external quality assurance (EQA) schemes. EQA organizations, like NordiQC, UK NEQAS and CAP, are independent organizations not associated with commercial suppliers. Their role is to promote the quality of immunohistochemistry (and in situ hybridization) by arranging external QA schemes for pathology laboratories. Similar EQA schemes are now available in many countries and regions around the world. The purpose of EQA schemes is to improve the quality of staining results in the participating laboratories; thus it is the individual labs that are being assessed. It is their choice of antibody, visualization system, instrumentation and protocol that is the basis for the EQA organization’s evaluation and feedback. A lab volunteers to participate in the assessment runs. Laboratories typically enroll for a year, during which they receive approximately 16 unstained tissue slides (NordiQC), or 7-8 different modules, where each module usually has two tissue slides (UK NEQAS), to stain using their own internal standard protocols for those markers designated by the QA organization. The labs return the stained slides to the QA organization for assessment, which is conducted by experts engaged by the organization. The labs receive either a “Passed” rating or “Not Passed” rating. Both NordiQC and UK NEQAS inform all participants of their individual scores and provide suggestions for protocol optimization when required. Both organizations present the anonymous results on their web sites, with statistics and best method for the particular marker.

CAP (College of American Pathologists) in the US, has a similar QA process, but requires only the return of stain results and interpretation, not the stained slides.

Some broad conclusions are possible:

- high-quality reagents are available, with highly sensitive detection methods, but they must be employed properly in controlled fashion, and currently often are not. Participation in EQA schemes can help laboratories improve the reproducibility;

- there is a pressing need for tissue-based IHC controls (or ‘reference standards’) (19, 25) that can be made available to all laboratories performing IHC assays, somewhat analogous to the international reference standards and calibrators that are available to clinical laboratories performing ELISA testing.

From this brief discussion it follows that to improve standardization to the point that all laboratories would carry out the IHC in identical fashion for every phase of the ‘Total Test’; it would require them to use the same fixative and fixation time (adjusted to tissue type), the same antigen retrieval process, the same primary antibodies and detection systems, with the same automated stainer and common controls. Clearly this perfect option will never happen, and we therefore must do what we can to reduce the consequences of the variables in the process.

Ultimately the overriding factor in effecting significant change must be to transform the mindset of pathologists, at least of the next generation, to the view that the end result of an IHC protocol is not just a ‘stain’, with intensity to be adjusted at the whim of the pathologist. Rather IHC is a precise immunoassay that is strictly quantifiable, and must be performed only with a degree of technical rigor and control that matches any other immunologically-based assay of like principle (namely ELISA). ELISA is a ‘gold standard’ method for quantitative assays in the clinical laboratory. ELISA reagents are purchased in prepared form, with all of the necessary reagents, defined protocols, and reference or calibration standards, for use with specified instrumentation. Ready-to-use reagents, coupled with proven detection systems, fixed and validated protocols, recommended controls and automation, represent an analogous pathway that could, if widely adopted with appropriate controls, lead to improved levels of reliability and performance for IHC.
References


