Antibody validation of immunohistochemistry for biomarker discovery: Recommendations of a consortium of academic and pharmaceutical based histopathology researchers

William J. Howat, Arthur Lewis, Phillipa Jones, Caroline Kampf, Fredrik Pontén, Chris M. van der Loos, Neil Gray, Chris Womack, Anthony Warford

Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge CB2 0RE, UK
MedImmune, Granta Park, Cambridge, UK
UCL Advanced Diagnostics, 21 University Street, London WC1E 6JJ, UK
Department of Immunology, Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden
Department of Pathology, Academic Medical Center, Amsterdam, The Netherlands
AstraZeneca Oncology Innovative Medicines, Macclesfield, UK
University of Westminster, London, UK

A R T I C L E   I N F O

Article history:
Received 20 September 2013
Revised 27 January 2014
Accepted 28 January 2014
Available online 11 February 2014

This article is dedicated in the memory of Neil Gray and Chris van der Loos, both of which were trusted colleagues and friends in the histopathology community. They will be sadly missed. Neil died on 31 July 2012 and Chris on 26 Nov 2013.

Keywords:
Antibody validation
Immunohistochemistry
Biomarker
Guidelines
Histopathology

A B S T R A C T

As biomarker discovery takes centre-stage, the role of immunohistochemistry within that process is increasing. At the same time, the number of antibodies being produced for “research use” continues to rise and it is important that antibodies to be used as biomarkers are validated for specificity and sensitivity before use. This guideline seeks to provide a stepwise approach for the validation of an antibody for immunohistochemical assays, reflecting the views of a consortium of academic and pharmaceutical based histopathology researchers. We propose that antibodies are placed into a tier system, level 1–3, based on evidence of their usage in immunohistochemistry, and that the degree of validation required is proportionate to their place on that tier.

1. Introduction

The term “biomarker”, in its broadest sense, defines any biological or physiological entity that is used to identify disease, guide targeted therapy or monitor for re-occurrence. In histopathology, immunohistochemistry (IHC) is routinely used for diagnosis [1] and with the advent of the patient selective cancer therapy trastuzamab for breast [2] and gastric cancer [3], or demonstration of C-kit for gastrointestinal stromal tumours targeted therapy [4], it is being used as a decision making tool to ascertain those patients who are most likely to benefit from treatment. Furthermore, the possibility of using antibodies to detect specific EGFR mutations as a guide for the administration of EGFR-targeted therapies in non-small cell lung cancer could result in IHC being used as a quick and cost effective replacement for the DNA sequencing based methods presently employed [5].

The development of an IHC biomarker can begin at the same time as the association of gene expression with a disease points toward development of a drug. Thus microarray data, next-generation sequencing and sometimes in situ hybridisation can provide the targets for biomarker selection and point to the need to either select or make an antibody to that target [6]. Whether ‘home grown’ or selected from existing commercial offerings it is of critical importance that the biomarker antibody is validated as specific for its target and of sufficient sensitivity to allow IHC demonstration over the required dynamic range demanded by the pathology it will be used to identify. The chief benefit of early validation is that the IHC based...
biomarker can be used with confidence during the drug development process to assist in understanding the target better, to segregate pathologies most likely to benefit from therapy and potentially to become the method by which this selection is made in the clinical setting. In the wider context of research pathology where IHC is frequently employed, thorough antibody validation will ensure that quality reagents are used. Regrettably, information supplied in many academic publications [7] and contained in commercial data sheets is not sufficient to allow confidence to be built into an antibody and on-going validation is required [8]. Thus, time needs to be expended by others to make good the information gap, a process that is inherently inefficient when an antibody is shown to be unacceptable for use. The purpose of this guideline, similar to guidelines published recently on tissue microarrays [9] and the efforts of the Clinical Laboratory Standards Institute for standardisation [10] is therefore to provide a tiered approach to the validation of an antibody for use as an IHC biomarker in formalin-fixed, paraffin-embeddded (FFPE) tissue and to promote this being undertaken before it is used as a biomarker of disease. These guidelines are equally applicable to the validation of an antibody for use in frozen tissue IHC and wholemount in situ staining protocols. A summary of the key features of these guidelines is contained in Table 1.

2. Steps to validation

2.1. Step 1: Understand target

It is vital before attempting any validation that a full literature review of the target is undertaken. Firstly, this will build a picture of when and where expression is to be expected and, should IHC have been attempted previously, can point to antibodies that could be evaluated. Secondly, where post-translational modifications or splice variants have been described, this information can be used to predict detection of multiple bands in Western blotting and thus antibodies that would have been rejected as “non-specific” will be kept. Databases such as OMIM [11], Uniprot [12] or Genecards [13] are particularly useful for gathering such information. Note, however, that online resources that are based on mRNA expression can provide spurious results, since the levels of protein and mRNA do not always correlate [14,15]. The biological relevance of the target is important, as this can give an indicator on the likely sub-cellular localisation of the target and as a consequence any non-specific interactions can be identified. For example, a transcribing factor is likely to have a nuclear localisation and therefore a cell surface staining pattern would be spurious.

2.2. Step 2: Identify cell and tissues

Control material is critical to the validation and can take several forms. Positive and negative cultured cells, identified through the literature search, can be used for Western blotting, flow cytometry (for membrane bound targets) and the preparation of FFPE cell blocks for IHC [16,17]. When selecting positive and negative control cell lines it is of great value to determine the expression of the required biomarker using more than one assay format e.g. confirming positive or negative biomarker expression of cells by flow cytometry before using as an IHC control. This builds confidence in the expression profile of control cell lines and forms part of the validation process. It is important when FFPE cell blocks are made, that cells are spun down lightly to retain cytology, fixed in the same buffers and times as tissue controls, suspended in agarose and processed in the normal manner to mimic the tissue that they are validating. The standardisation of this process is equally essential, since a variety of fixation regimes and processing methods are referred to in the literature. Once individually processed, the creation of a cell line microarray (CMA) [18,19] can assist not only in determining reaction across multiple cell lines, but also as a quality control check of the IHC once the biomarker antibody has been validated [20]. Using cell lines in this manner has already been validated in the clinical setting [21].

The use of cell culture lines can also prove beneficial particularly where they can be manipulated by transfection to introduce different ‘dose’ levels of the target in otherwise weakly positive or negative cell lines. As transfection efficiency rarely reaches 100%, a proportion of the cells should remain negative or weakly stained for the target in question which can be useful in differentiating IHC signal from background noise. This also remains the case where RNA is used to knockdown positive cell lines. Indeed, a recent confocal study on cell lines has demonstrated that in 765 proteins studied using 75 antibodies that siRNA silencing can be effective in 80% of cases [22].

Whilst cell lines can provide an indicator of the expression, particularly in comparison to non-IHC methods, the use of positive/negative control tissue is essential for the full validation of the antibody and evaluation of non-specific binding to other tissue components. The literature and online reviews should point to potential positive and negative tissues. However, as expression levels are often modulated in disease then it may be important to include a range of pathologies and, preferably, matched normal tissue. Prior to selecting positive material, the quality of the tissues needs to be assessed and it is therefore recommended that a panel of

<table>
<thead>
<tr>
<th>Step</th>
<th>Special considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understand target</td>
<td>Full literature review&lt;br&gt;Note the biological relevance and expected sub-cellular localisation</td>
</tr>
<tr>
<td>Identify cells and tissues</td>
<td>Identify or create positive/negative cell lines&lt;br&gt;Identify positive/negative control tissues&lt;br&gt;Check the quality of the control material with standardised antibodies and ensure it matches the quality of the test material</td>
</tr>
<tr>
<td>Choose an appropriate IHC method</td>
<td>Use a commercially prepared and validated kit&lt;br&gt;Test multiple retrieval conditions to optimise the staining&lt;br&gt;Consider the use of automation</td>
</tr>
<tr>
<td>Identify level of validation required</td>
<td>Identify which tier of validation is appropriate&lt;br&gt;Tier 1: Well known antibody with high quality literature evidence&lt;br&gt;Tier 2: Well known antibody used in an alternative species or unvalidated tissue&lt;br&gt;Tier 3: Unknown antibody. Inconsistent/No literature evidence</td>
</tr>
<tr>
<td>Control of IHC</td>
<td>Use the same controls that were used for validation when performing test samples&lt;br&gt;Test antibody in at least one other non-IHC method&lt;br&gt;Use negative controls, such as omission of primary antibody or isotype-matched controls to identify any background staining</td>
</tr>
<tr>
<td>Publication of results</td>
<td>Include appropriate control material either within publication or as supplementary material&lt;br&gt;MISFISHIE guidelines should be used</td>
</tr>
</tbody>
</table>
characterised generic antibodies, preferably in vitro diagnostic reagents, are run on the tissue to validate it. For example, desmin for muscle (e.g. Clone DE-R-II), cytokeratin for epithelial tissue (e.g. Clone AE1/3), neurofilament for neural tissue (e.g. 2F11).

Some targets, such as phosphoproteins, require special consideration in tissue choice as ischaemia can dramatically reduce the level of target available for IHC demonstration [23]. Due to clinical constraints, there is little opportunity to control warm ischaemia but samples should be immersed in formalin (or otherwise preserved within 15–30 min of removal from the patient). A pan-phosphotyrosine antibody can have utility in determining whether a tissue sample is unlikely to demonstrate phospho-epitopes [24]. Holding tissue after fixation in ethanol does not appear to have any adverse effects on the levels of phosphoproteins demonstrable by IHC [25], however, it is recommended that a standardised time in ethanol is used per study to ensure comparable results. These factors, of course, can put constraints of the application of this type of antibody once the validation process has been completed.

As with cell lines the preparation of TMAs is very useful as this will maximise IHC data whilst ensuring that the technical parameters are uniformly applied to the tissue set. It has been reported that some antigens, i.e., CD3 [26], ER [27] and pSMA2 (WH personal observation), deteriorate in cut sections which has recently been proposed to be due to presence of water within inadequately processed tissue specimens [28]. To avoid this becoming a complicating factor in the validation of a new antibody, sections should be cut and used within a standard period of time. The practice of cutting sections and coating with paraffin wax alone to minimise surface oxidation/hydrolysis has not been routinely shown to be of benefit [27,29], however, the storage of tissue microarray sections at −20 °C without coating in wax has been favourably tested to room temperature storage over a 6 month period (Kampf, unpublished observations).

2.3. Step 3: Choose an appropriate IHC method

The expected expression level of the target will often determine the choice of detection method. For example, the Labelled Streptavidin Biotin (LSAB) technique may not have the sensitivity for some antigenic expression, where use of polymer systems or tyramine amplification may be required in its stead [30,31]. The use of esoteric detection methods should be avoided as these may lead to problems of reproducibility. Care should also be taken to approach variation of the technique systematically, to employ the same reagent batches throughout the validation and to record staining results using a standardised reporting template, such as the ‘H’ score system [32], so that run-to-run comparisons can be made.

Unless the literature review indicates otherwise, the comparison of pH6 and pH9 heat induced epitope retrieval (HIER) buffers together with incubation of the antibody at two dilutions/concentrations should provide an initial indication of which combination is worth pursuing further. The addition of an enzyme digestion retrieval at this stage may reveal antigens that are particularly sensitive to protease digestion, such as alpha smooth muscle actin, clone 1A4 [33]. If these assessments fail to provide adequate staining then consideration should be given to using a combination of protease digestions and HIER or no retrieval at all. In all cases, the staining pattern received should be compared to that described in the literature for alignments to tissue positivity, cellular positivity and sub-cellular localisations, as a variation of the expected staining pattern more often indicates a non-specific binding of the antibody rather than an interesting scientific observation. Other variations that can be tested include the time and temperature of HIER, primary antibody incubation and range of antibody dilution/concentration. Alternative fixatives can be investigated but can only be applied to prospective collections, as the majority of tissue collections in both research and clinical arena’s utilise neutral buffered formaldehyde.

The above principles apply whether performing manual or automated staining. The use of semi- or fully-automated staining systems has revolutionised immunohistochemistry particularly in the clinical arena, where in vitro diagnostic antibodies and staining kits are employed to provide fast and reproducible staining which is computer controlled and monitored. Thus, automated IHC has significant advantages and can provide the researcher with reproducibility, sensitivity, a reduction in error rate and improvements in slide tracking and monitoring. However, the cost of such systems can be prohibitive to smaller research laboratories running lower throughput assays and there is no evidence that an automated assay outperforms a well validated manual assay in this setting.

2.4. Step 4: Identify level of validation required

We recommend the following 3 tiers of validation:

Level 1: Where an antibody is to be applied to tissue for which there is reliable IHC literature and it has become well established and trusted; an example are antibodies that recognise Ki67 [34], then validation can be as straightforward as reproducing the expected result on positive and negative tissues to gain an appropriate signal/noise ratio.

Level 2: Where an antibody has previous cited literature for IHC that is reliable, but where the tissue or preparation is different to that previously described, then validation should be undertaken using positive control material as detailed in Step 2 (above). For example, the application of HER2 in gastric cancer, where it was originally validated in breast cancer, would fall under this category. The resulting IHC should be carefully compared for consistency of staining and, where necessary, the dilution/concentration of the antibody adjusted to give optimal signal/noise ratio in the new samples. Where an inconsistent pattern appears at this point, the validation should be taken to Level 3, as it may be that the previous published data is unreliable [35].

Level 3: This level of validation is required where little or no reliable IHC data is available. The most extreme example of this would be a home-grown antibody, however, an antibody where there is no evidence on the datasheet that it has been tested by immunohistochemistry would also fall under this category. Similarly, with experience, antibodies designated by the supplier as appropriate for IHC, may also fall into this category where insufficient data is provided to demonstrate this, or where that data is inconsistent with the knowledge gained under Step 1 above.

In this instance validation should proceed in a stepwise fashion to build confidence in the performance of an antibody. Where several antibodies are available then these can be taken through the process together and selectively discarded if they fail at a validation step. The use of “sibling antibodies”, where multiple antibodies are prepared to the same target, as utilised by the Human Protein Atlas and Atlas of Protein Expression projects [36,37], can add confidence in the data quality and thus may be considered as a Level 2 validation, but only where both antibodies provide the same staining pattern.

Before running the antibody in IHC, it must be tested on at least one other non-IHC method. As a minimum, this should be an in-house Western blot, where the protein is resolved from a cell or tissue lysate or from formalin-fixed tissue sections [38]. The use of Western blots generated against recombinant protein alone is not recommended as it unlikely to demonstrate the full range of potential cross-reactivities. Flow cytometry against panels of known positive cell lines can demonstrate specificity and a degree of cross-reactivity of cell membrane antibody targets. Ultimately,
the antibody may be used in an immunoprecipitation experiment, cut out and resolved by mass spectrometry for specificity \[39,40\]. Before running the antibody on tissues, it must be tested on putative positive and negative control material, since determination of specificity by Western blot does not guarantee specificity after the target has been subjected to formalin fixation and antigen retrieval. The positive tissues should establish that the correct cell type, compartment and staining intensity is present, whilst negative tissues are valuable in confirming the non-reactivity of the antibody and adjusting signal/noise ratio if background staining is encountered. When used in combination with positive or negative cell lines, either inherent or induced, the specificity of the antibody will be confirmed. Once an antibody has passed through these validation steps it should be used on TMA preparations of relevant tissues to confirm its sensitivity and specificity on a sample set representative of its ultimate use.

2.5. Step 5: Control of IHC

Once an antibody has been successfully validated, run controls must be used to confirm that acceptable results are generated. In each IHC run both positive and negative controls should be included.

When possible the positive control should match the tissue under investigation and should contain a moderate level of target. Thus, if conditions of an IHC run are sub-optimal the intensity of staining in the positive control will be noticeable and the investigation of potential reasons for this can begin. For clinical antibody based biomarkers, such as HER2 \[41,42\], the incorporation of a range of control tissues or FFPE cell blocks preparations can be obligatory to confirm that sensitivity levels are being met.

Negative controls provide information on aberrant staining due to the presence of endogenous substances, such as endogenous biotin or peroxidase, or other non-specific staining due to binding of the primary and/or secondary reagents. Thus they control both the staining method as well as the validated antibody. Omission of the primary antibody only and replacement with primary antibody diluent, controls only the staining method. For control of a monoclonal primary antibody, an isotype matched control fraction should be used whereas controlling for affinity-purified polyclonal antibodies, an affinity-purified, species-matched polyclonal or incubation in non-immune serum should reveal general non-specific staining. However, it should be noted that the use of the latter control can itself give rise to this type of staining because of the milieu of proteins present in its composition. Ultimately, an isotype matched antibody for a target that is absent from the tissue, for example GFP or BrdU, can be used, but these are not always available.

The employment of blocking peptides in an IHC reaction should be carefully considered. Whilst suppression of signal is evidence that the primary antibody is interacting with the tissue this cannot be used as prima facie evidence of specificity as cross-reacting substances will also be suppressed \[43\]. Abolition of IHC signal by a phosphatase treatment step can be used to demonstrate in the validity of phospho-specific antibodies.

2.6. Step 6: Publication of results

To avoid the need for revalidation of already published material or to enable validation to be taken further, all IHC publications should include information that clearly states the source of antibody employed and the conditions under which it was used; Table 2 sets out an example of this. At this time, only the Journal of Comparative Neurology \[44\], and more recently Endocrinology \[45\], specifically request their authors to state their antibody’s specificity and validations. Indeed, the Journal of Comparative Neurology has gone a step further to publish a free online database of all antibodies submitted to the journal since 2006 \[46\]. To this end it is strongly recommended that the MISFISHIE guidelines for reporting in situ hybridisation and IHC is used \[7\] and it should be noted that these guidelines also include requirements for standardised scoring systems and imaging. Furthermore, reporting of the conditions on an open access site following publication, such as AntibodyPedia \[47,48\] can only increase the use of antibodies for biomarker validation and improve the quality of the publications brought from their use.

3. Summary recommendations

IHC has become the ‘special stain’ of histopathology. It has an important diagnostic role, and is increasingly used as a biomarker to phenotype tumours leading to targeted drug therapy. Unlike some tinctorial staining methods, the techniques that are employed in IHC are relatively few and reproducible when undertaken in a careful manner. However, whilst antibodies used as biomarkers for targeted drug therapy are, of necessity, validated and approved by regulatory bodies, the evolution of many antibodies that are now accepted for research use in IHC has been through a path akin to survival of the fittest. There are thousands of biomarker candidate antibodies available that could be used for IHC, but few carry with them any evidence of rigorous validation. Furthermore, many antibodies have been used in publications without due consideration of the necessity of validation or indeed the recording of methods used in sufficient detail to allow the work to be reproduced \[7\]. This in turn has led to false starts in IHC based biomarker identification and wasted research effort. With these realities in mind we have provided guidelines as the basis of best practice for biomarker antibody validation. Our recommendations are based on a stepwise approach by which confidence is gradually built in an antibody and end with the information acquired during the validation process being shared with the scientific community to avoid the need for its repetition. Adoption of the guidelines should assist biomarker development and also benefit the IHC community as a whole, enabling the identification of antibodies that will make both diagnostic and clinical differences.

Acknowledgments

The authors would like to acknowledge the University of Cambridge, Cancer Research UK and Hutchison Whampoa Ltd.

References
