



Published in final edited form as:

Biotechniques. 2010 March ; 48(3): 197–209. doi:10.2144/000113382.

Antibody validation

Jennifer Bordeaux, Allison W. Welsh, Seema Agarwal, Elizabeth Killiam, Maria T. Baquero, Jason A. Hanna, Valsamo K. Anagnostou, and David L. Rimm

Department of Pathology, Yale University School of Medicine, New Haven, CT, USA

Abstract

Antibodies are among the most frequently used tools in basic science research and in clinical assays, but there are no universally accepted guidelines or standardized methods for determining the validity of these reagents. Furthermore, for commercially available antibodies, it is clear that what is on the label does not necessarily correspond to what is in the tube. To validate an antibody, it must be shown to be specific, selective, and reproducible in the context for which it is to be used. In this review, we highlight the common pitfalls when working with antibodies, common practices for validating antibodies, and levels of commercial antibody validation for seven vendors. Finally, we share our algorithm for antibody validation for immunohistochemistry and quantitative immunofluorescence.

Keywords

antibody; validation; immunohistochemistry; immunofluorescence

Introduction

Antibodies are among the most commonly used research tools, routinely used for Western blot (WB), immunoprecipitation (IP), enzyme-linked immunosorbent assays (ELISA), quantitative immunofluorescence (QIF), and immunohistochemistry (IHC). They are also important tools in clinical management with extensive use in both laboratory medicine (ELISA assays and flow cytometry) and anatomic pathology (IHC). In anatomic pathology, IHC serves as a diagnostic, prognostic, and predictive method and IHC readings directly influence the management of patients in the clinical setting. For example, the assessment of estrogen receptor α (ER- α), and human epidermal growth factor receptor 2 (HER2) by IHC in breast cancer patients is the definitive test to determine whether or not a patient will receive therapies that can cost as much as \$100,000 per year. Thus, in the clinic, as well as in the research laboratory, careful accurate validation of antibody reagents is critical for correct results.

The influence of antibody-based tests on clinical decisions has led to a number of publications that have highlighted the unmet need for standardization of such assays and development of antibody validation guidelines (1–8). Although many groups have enunciated the need, there are no universally accepted guidelines for best practice in antibody-based tests. There are a number of books on the topics by world leaders such as Clive Taylor and David Dabbs, and recently, an ad hoc group published a set of

Address correspondence to David L. Rimm, Department of Pathology, Yale University School of Medicine, 310 Cedar St., PO Box 208023, New Haven, CT, 06520 USA. david.rimm@yale.edu.

Competing interests

The authors declare no competing interests.

“recommendations” (2). However, these works focus on the clinical aspects of IHC, often using subjective criteria and often not taking advantage of recent scientific advances that allow more quantitative evaluation of antibodies. Conversely, there are other groups that have done biologically rigorous evaluation of antibodies using surface plasmon resonance (9) and even X-ray crystallization of antibodies bound to their antigens (10), methods that are unachievable in a routine research or clinical setting. The wide range of rigor and methodology in what is used for validation is probably responsible for a lack of consensus on a single method for antibody validation. Here we present an overview of antibody validation approaches and the pitfalls associated with the failures of validation. This work specifically focuses on assessment of prognostic and predictive cancer-related biomarkers on formalin-fixed paraffin embedded (FFPE) tissue.

What is antibody validation?

The FDA defines validation as “the process of demonstrating, through the use of specific laboratory investigations, that the performance characteristics of an analytical method are suitable for its intended analytical use” (www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070107.pdf). For antibodies, one must demonstrate that they are specific, selective, and reproducible in the context for which they are used. When it comes to IHC, standardization can be quite challenging due to the number of pre-analytical, analytical, and post-analytical factors known to influence staining in FFPE tissue. Variable time to fixation, inadequate fixation period, differences in fixative used, and tissue processing can all affect tissue antigenicity (5,11). Antibody clone and dilution, antigen retrieval, detection system, and interpretation of results using different cutoff points are also important variables that regulate IHC measurements (3,12) (V.K.A., unpublished data). Here we focus on analytical factors and highlight the importance of proper antibody validation, especially for IHC or QIF use.

Common pitfalls in antibody studies

Nonspecific antibodies

A recent editorial by Michel et al. (13) emphasizes the lack of target specificity for 49 antibodies against 19 subtypes of G protein-coupled receptors calling for more stringent antibody validation criteria. Examples highlighted by the authors included double-knockout mice lacking the M₂ and M₃ subtypes of muscarinic receptors still staining positive for M₂ and M₃ receptor antibodies (14), and triple-knockout mice for the three α_1 -adrenoceptor subtypes demonstrating staining patterns similar to wild-type mice (15).

Determining the specificity of an antibody is in part dependent on the type of the immunogens: synthetic peptides or purified proteins. Synthetic peptides provide the advantage of knowing the amino acid sequence to which the antibody binds; however, these peptides do not necessarily recapitulate the 3-D structure or post-translational modifications of the native protein (16). As a result, antibodies generated against a synthetic peptide may not work well when a protein is in its native conformation with intact 3-D structure. Such antibodies may not be useful for IP or IHC experiments, but may bind the protein of interest after it is fully denatured when running SDS WB. The opposite could also be the case, especially if the immunogen was the purified protein, where the antibody works well for proteins in their native conformation, but not when denatured. Thus WB cannot be an absolute standardization for antibody binding in IHC or other assays where the antigen is in its native conformation.

The native versus denatured confirmation is further complicated by methods used to fix tissue. Epitopes that are not exposed in the native proteins can be exposed in fixed tissue and

vice versa, even though they may not be truly denatured. Thus, an antibody could recognize one epitope in fresh tissue, but when applied to fixed tissue recognize another epitope (17,18). A representative example is an epitope on BCL-2 (41–54 amino acids) that is exposed when BCL-2 is present in the cytoplasm but is inaccessible in the nuclear compartment most likely due to interaction with other proteins (19). However, when BCL-2 is phosphorylated at sites close to the epitope, the epitope becomes available, as seen when the protein is extracted from cells or denatured by SDS (20).

The issue of epitope specificity is further complicated by the choice of monoclonal versus polyclonal antibodies. Polyclonal antibodies represent a pool of antibodies against the immunogen and typically show a higher probability for detection in a range of different conditions, whereas a monoclonal antibody is more likely to work in only one set of conditions. Polyclonal antibodies might nevertheless contain superior affinity clones. Perhaps the best example of this is the FDA-approved Herceptest antibody from Dako, which is the industry standard for HER2 testing in breast cancer (21). Monoclonal antibodies are more pure, since they are produced from single clones of fused cells producing immune globulins. However, these clones may be grown in host animals where the ascites fluid containing the secreted antibody is collected (16,18); therefore, these antibody preparations may be contaminated by antibodies other than the monoclonal antibody of interest. A study by Spicer et al. demonstrated that seven of 20 monoclonal antibody preparations (35%) they analyzed had staining patterns localized to the Golgi cisternae unrelated to the antigenic specificity of the antibody, and that five of these cross-reactive antibodies failed to even stain the antigen of interest (22). Antibody validation efforts in our lab have shown similar lack of specificity, even in monoclonal antibodies. Figure 1 provides two examples of nonspecific antibodies tested in our lab. In Figure 1A, a mouse polyclonal antibody against HoxA1 (B01P; Abnova, Taipei City, Taiwan) was used to probe lysates from ten cell lines by WB. The expected molecular weight is 37 kD and a band of this size was seen in the CaCo2 lysates (denoted by the arrow). All of the lysates show several bands above and below this expected molecular weight at a lower signal level, such that if this antibody were validated only against a cell line transfected to overexpress HoxA1, these bands would have been missed. This level of noise as seen by WB raises concerns about nonspecificity. Figure 1B provides a representative example of the staining pattern seen with this HoxA1 antibody on breast carcinoma tissue. The predominately cytoplasmic staining for a homeobox transcription factor indicates that this antibody should not be used for IHC or IF methods. Figure 1C provides an example of a monoclonal antibody for phospho-4EBP1 (rabbit MAb, clone 236B4; Cell Signaling Technology, Beverly, MA, USA) showing a band of the expected molecular weight (denoted by the arrow) as well as numerous additional bands at higher molecular weights. Figure 1D is a representative example of the nuclear staining observed when using this antibody for IHC on FFPE lung carcinoma tissue when the expected localization for p-4EBP1 is cytoplasmic.

Non-reproducible antibodies

Poor correlations between antibody lots are cause for concern, as demonstrated by a recent study from our lab on antibodies against the Met tyrosine kinase receptor (23). Strikingly, two different lots of the monoclonal 3D4 Met antibody stained on an array of 688 breast cancer cases showed opposite staining patterns—one nuclear and one membranous and cytoplasmic—with a regression between the two lots having an R^2 value of 0.038 (23). Another example of non-reproducible antibodies was demonstrated by Grimsey et al. (2008) with antibodies supposedly targeting the cannabinoid CB₁ receptor. The authors tested multiple lots of antibodies from various commercial sources by both WB and IF against HEK cells transfected with HA-tagged CB₁ receptor. Only two of the six antibodies tested showed specific membranous staining that co-localized with detection using an anti-HA

antibody. Additionally, WB analysis on HEK cell lysates demonstrated that the antibodies with poor IF specificity detected proteins of incorrect molecular weight, no proteins at all, or proteins present in wild-type HEK cells which do not even express the CB₁ receptor (24).

Some examples of non-reproducibility can be very subtle. Figure 2 shows two examples of what appears to be specific staining with expected localization. However, staining with the same conditions on a serial section of the TMA (representing the same patient tissues) shows an extremely poor correlation. VEGF is a secreted growth factor and as such, we expected cytoplasmic staining on patient tissues with the VG-1 clone (Figure 2A). Staining on a serial section of the same tissue microarray with the same lot of antibody was not reproducible, with an R^2 value of 0.016 (Figure 2B).

Current practice in antibody validation

Antibody specificity

WB is widely used to determine an antibody's specificity and is an appropriate first validation step if the antibody recognizes the denatured antigen. The first indication that the antibody is specific for the selected target would be observing a single band at the known molecular weight for the target. Presence of multiple bands or bands not at the proper molecular weight could represent the same target at different post-translational modification status, breakdown products, or splice variants. However, such observations should raise concerns for using this antibody for further experiments. Major et al. (2006) published their database AbMiner as a resource for information including the immunogen, vendor, and antigen on over 600 commercially available monoclonal antibodies that the group validated via WB against pooled cell lysates from each of the NCI-60 cell lines. An antibody was considered validated if it produced a band (or bands) of the expected molecular weight(s) for the target protein (25).

While this type of antibody validation is a useful first step, it only guarantees that a given antibody will provide accurate results for WB analysis. If the goal is to use the antibody for IHC or IF, then the user must demonstrate that the antibody is also able to specifically recognize its target when used for those applications. Antibody-pedia is also an online portal for sharing antibody validation data for WB, IHC, and IF. Only antibodies commercially available are submitted, and in addition to a validation score, the inclusion of original experimental data is required. (26).

Antibody specificity has also been evaluated by using blocking peptides especially for IHC (27). These peptides are the sequences used to generate the antibody and are incubated with the antibody in great excess. The antibody with and without the blocking peptide is then used to stain tissue known to express the target of interest. If the antibody is specific, the addition of the blocking peptide will result in loss of staining on the tissue. An example using this technique was recently published (27) for the validation of phospho-specific antibodies for ER α . Although this method demonstrates that the antibody is specific for the immunogen from which it was generated, it does not prove selectivity of the antibody since off-target binding activity of the antibody will also be inhibited by pre-adsorption with the blocking peptide. So while blocking peptides can prove that an antibody is bad, when nonspecific staining is seen in the presence of the peptide, they cannot prove that an antibody is good. Blocking peptides had been used with a number of the G protein-coupled receptor antibodies and have been found not to be selective upon more stringent validation (13). Thus, we do not typically include blocking peptides as part of our antibody validation process.

The key to proving antibody specificity is often the correct use of controls. A negative control, including no primary antibody in an IHC or QIF experiment, is valuable but insufficient. A better negative control is a cell line or tissue that is known not to express the protein of interest. Knockout cells thus provide the best negative controls. Similarly, non-expressing cells, transfected with the protein of interest, provide the best positive controls. Since these reagents are often beyond the reach of many labs, there are other approaches that can be used to obtain comparable control results. Often there are readily available cell lines that have been biologically proven not to express a specific protein of interest. For example, H1650 cells are PTEN-null and thereby make a good negative control for PTEN antibodies. Similarly, overexpression can be valuable as a positive control. For example, A431 cells overexpress wild-type EGFR and can be used as a positive control for anti-EGFR antibodies. An alternative to knockout cell controls is siRNA or shRNA knockdown controls, as will be discussed with our Rimm Lab Algorithm of antibody validation.

Finally, another type of quality control slide has been developed by Sompuram et al. that couples peptides to glass slides that mimic the epitope of the native antigen and are therefore specific controls for a given monoclonal antibody (28). The drawback to this type of control is that it only works for monoclonal antibodies, as the antibody must interact with a known epitope in order to generate the appropriate peptide controls.

Antibody reproducibility

An important criterion for validation and standardization is antibody reproducibility. This means staining of the same antibody over time with different lots on different days, as well as comparison of a new antibody to either a previously validated antibody or to a second independent means of measuring the target that would yield similar results (29). There are not many studies on reproducibility of antibodies, since this is often assumed when using new lots of the same antibody or new aliquots of an antibody used previously by the same lab. Our lab published one such example of this type of evaluation (23), and a similar work to show assay reproducibility has been done for HER2 by Gustavson and colleagues (30).

A second issue related to reproducibility is when a new reagent is introduced and there is a desire to compare the new antibody to previously validated standards. This sort of study is seen more commonly in the literature. For example, van der Vegt et al. compared HER2 IHC staining of 283 breast adenocarcinomas on tissue microarrays with the 4B5 rabbit monoclonal antibody to the previously established CB11 mouse monoclonal antibody. Additionally, both antibodies were also compared to each other and to the corresponding FISH scores for each case. The authors found no significant differences in sensitivity, specificity, or predictive values between the two antibodies, validating that 4B5 could be used to assess HER2 expression (31). A second example by Zafrani et al. compared antibodies for ER and PR with IHC to biochemical determination using enzyme immunoassay and found significant links between the two methods (32). A similar study by Sasano et al. validated the use of an aromatase monoclonal antibody by correlating the IHC score with biochemical activity as determined by product isolation assays (33). Sometimes this is done in an attempt to prove superiority of a new reagent. In those cases, investigators must compare the reagents in the setting of intended use with outcome or response to therapy, not the previous assay, as the criterion standard. A good example of this is the work of Cheang et al. that showed the ER antibody SP-1 to be more prognostic and predictive than the current standard 1D5 (34), although this conclusion was later disputed by Brock and colleagues (35).

Vendor-based validation of commercial antibodies

It is estimated that there are over 180 antibody companies that produce over 350,000 antibodies for the research and clinical markets (www.antibodyre-source.com/onlinecomp.html). So it is not surprising that there are commercially available antibodies against a huge collection of target proteins. When one buys a kit to isolate DNA, it is generally safe to assume it will not end up purifying protein. However, the same is not true for antibodies. Just because the manufacturer claims an antibody is specific for protein Z does not necessarily mean that it will only bind protein Z and not a range of other proteins (29,36). In 2005, Ramos-Vara noted that the information about antibodies targeting the same antigen varied significantly depending on the manufacturer (16). In fact, it is now clear that the responsibility for proof of specificity is with the purchaser, not the vendor. Different vendors provide different levels of validation, depending on their approach to the balance between making a profit and providing high quality. For this review, we sought to compare the level of validation and amount of information provided by seven different companies (referred to herein simply as Companies 1–7) to the randomly selected molecule AMPK β_1 . The companies are not revealed since only 7 of over 180 were chosen at random; furthermore, the level of validation at some companies varies based on the product. When available, we compared the validation protocol on the datasheet for phospho-specific antibodies, as these require another level of specificity and validation (37,38).

In general, we found three levels of validation. The least validation was observed for Company 1's phospho-AMPK β_1 polyclonal antibody. No information was readily available on its web site or datasheet describing any level of antibody validation criteria. The datasheet itself contained minimal information about the antibody including a brief background on the target with corresponding references, recommended applications and starting dilutions for WB, IF, and ELISA. However, no examples of successful use in any of these applications are provided and none of the references provided had utilized this antibody. The datasheet also includes the animal host and immunogen source as a synthetic peptide, but does not include exact peptide sequence, although the peptide is available for purchase for use as a blocking peptide. The datasheet also cautions that this antibody "may" cross-react with correspondingly phosphorylated AMPK β_2 .

A moderate level of validation was observed for Companies 2–5. These companies also did not provide any in-depth descriptions of antibody validation procedures. Datasheets all included background on the target, as well as information on the immunogen. Three out of four companies included the complete sequence used and the fourth identified the region surrounding the phosphorylation site as the sequence used. These companies also provided recommended applications with starting dilutions, and all provided at least one example of the antibody successfully identifying its target in one of the recommended applications. WBs—either with transfected cells expressing the target or pre-incubation of the antibody with blocking peptide—were the most commonly shown antibody validation examples.

The highest levels of validation were seen for Companies 6 and 7. Company 6 describes its validation procedures for phospho-specific antibodies to include WB analysis in (i) multiple cell lines, (ii) peptide and phospho-peptide competition experiments, and (iii) analysis of site-directed mutants. It also strived to demonstrate lot-to-lot consistency aided by combination of two or more pre-qualified individual lots to create each batch to minimize batch-to-batch variability. Company 7 describes their stringent validation protocol for all antibodies to include testing of each antibody by WB, IP, IHC, IF, flow cytometry, and ELISA. It verifies the specificity and reproducibility of antibodies by using appropriate kinase-specific activators or inhibitors when available; testing against a large panel of cell lines with known target expression; phosphatase treatment for phospho-specificity;

comparison to isotype control; verification in transfected cells, knockout cells, and siRNA-treated cells; utilizing blocking peptides to eliminate all signal; verifying correct subcellular localization or treatment-induced translocation; comparing new antibody lots to previous lots; providing optimal dilutions and buffers, as well as specifying both positive and negative control cell lines. For IHC, Company 7's antibodies are tested on paraffin-embedded cell pellets including cell pellets created after the cell lines are subjected to treatments known to induce signaling changes or treated with siRNA to block expression of the target. Tissue is treated with phosphatase to additionally test phospho-specificity on FFPE tissues. Xenografts with cell lines of known target expression or treatments to modulate expression are paraffin-embedded and then stained.

The datasheets for these companies include everything seen with Companies 2–5 plus additional examples of the successful use of their antibody. For example, Company 6 includes representative data for all validated applications with recommended antibody concentrations and includes flow cytometry, immunocytochemistry on a positive control cell line (including incubation with either the phospho-peptide or the non-phospho-peptide demonstrating signal absence only with the phospho-peptide), IHC of a positive control tissue, and a WB of a positive control cell line lysate also including detection after pre-incubation with the phospho- or non-phospho-specific peptides demonstrating a single band of the expected molecular weight with signal absent only after incubation with the phospho-specific peptide.

Antibodies purchased from Company 1 (which provided the minimal amount of information and no examples of successful use of the product) would require extensive validation by a researcher to demonstrate that the results describe only the target of interest; the company provides little in the way of assurances that it will work. Companies 2–5 all provide more information on their respective antibodies and include at least one example of the product successfully being used for at least one of their recommended applications with corresponding positive controls. While this is clearly preferable to Company 1's approach, these antibodies would still need to be validated by the researcher for target specificity in the application of interest and for lot-to-lot reproducibility. Company 6 describes the validation steps it uses to determine that its phospho-specific antibodies are both specific and reproducible. Furthermore, for each recommended application, an example of successful use is provided with corresponding inhibition by phospho-specific blocking peptides. Company 7 provides what we consider the gold standard for antibody validation. The extensive in-house testing described provides a high level of confidence that the antibodies it provides will work for all the applications that are recommended for the particular antibody. It also provides a greater level of confidence that results obtained will be specific for the described target and will be reproducible among antibody lots. Even though Companies 6 and 7 provide extensive validation, the researcher is still obliged to confirm that the product gives specific and reproducible results in the cell lines or tissues of interest in the lab. Indeed, even the best companies cannot control what happens after the product leaves their door. Issues during shipping, inappropriate storage on or after arrival in the laboratory, and antibody contamination during usage are all potential sources of error in antibody-based testing. Thus vigilant and comprehensive controls should be done with each assay.

The Rimm Lab Algorithm for antibody validation for IHC/QIF

There are no uniform or enforceable standards for antibody validation. Unlike drugs—whose sale is prohibited without FDA approval—there is no federal agency governing what can be sold into the antibody-based assay market. In the future, we may see further FDA clearance of antibodies or more rigorous labeling and regulation of reagents to be used in clinical testing. However, to date, there is no universal standard. We therefore present our

lab's approach (the Rimm Lab Algorithm) to the validation issue. Our approach is not sanctioned or approved by any governing body or by any trade association, and thus we are not so bold as to call this a "recommendation." Instead, we provide what we feel is a compromise between rigor and lab economics that results in a level of evidence sufficient for data dissemination. Our algorithm for antibody validation (Figure 3) is especially focused for the end use application of IHC or QIF on paraffin-embedded tissues, but could be equally valid or modified for other antibody-based assays.

Our first line of evidence that an antibody is specific for the target of interest is done by WB. A variety of cell line lysates (or tissue homogenates) are selected, ideally with known levels of target expression such that both positive and negative cell lines are analyzed. Practically speaking, we often have no idea of target levels and select cell line series at random. Lines are often selected in such a way that we suspect some of the lines to be completely negative for the target protein (e.g., using fibroblast lines for epithelial targets). When true negative lines are not available, or when levels of target protein are highly dependent on growth conditions, we produce lysates of cell lines where the target protein has been knocked down using RNAi or lysates from cell lines typically null for the target that have been transfected for overexpression constructs. In a recent example, we have used inducible overexpression (A.W.W., unpublished data). An example of using siRNA to aid in validating an antibody is shown in Figure 4 for Stathmin (rabbit monoclonal; Epitomics, Burlingame, CA, USA), a microtubule-destabilizing protein. Lysates from BT-20 cells either mock-transfected or transfected with scrambled siRNA control or siRNA specific for Stathmin were analyzed by WB demonstrating loss of Stathmin expression (Figure 4A). BT-20 cells transfected with either scrambled or Stathmin siRNA were also fixed and decreased Stathmin expression was visualized with IF (Figure 4B). Additionally, when known target activators and inhibitors are available, lysates from treated cells can also be included when validating antibodies specific for a particular state (e.g., phosphorylation). Identification of a single band (or multiple bands if more than one isoform of the protein is expected to interact with the antibody) of the correct molecular weight only in cell lines expressing the target encourages further validation for use in IHC/QIF. An antibody demonstrating no binding on WB may still be specific for its intended target when in its native conformation and an IP experiment can be the next step in determining the specificity of the antibody if the goal is for use with IHC/QIF.

The second step in the validation of an antibody for IHC/QIF is to titer the antibody on a tissue microarray (TMA) comprised of FFPE cell line pellets corresponding to the identical cell lines used to first validate by WB in addition to patient tissue expected to express the target (for more information on the usefulness of TMAs, see References 39 and 40). A good antibody will have the following characteristics: (i) it will stain only the cell pellets expressing the target, (ii) the level of staining will decrease with increasing dilutions of the antibody, and (iii) it will demonstrate an expression pattern that is consistent with biological and mechanistic data in the published literature. The target expression seen with these first two steps should be quantified—we use ImageJ for quantification of WBs and A3UA for measuring the level of protein expression in each TMA spot—and then the results should show a strong correlation to each other. An example of an antibody meeting these conditions is shown in Figure 5 for ER- α (1D5 clone; Dako, Glostrup, Denmark). As expected from published literature, the staining on patient breast tissue is predominately nuclear (Figure 5A). Breast cell lines with varying ER expression levels (Figure 5B), as well as a calibration curve of recombinant ER protein, were analyzed by WB and the expression of ER protein (pg per μ g of loaded cell lysate protein) was plotted against the expression level of ER as determined by IHC and quantified by A3UA demonstrating a strong correlation (R^2 value of 0.91) between the two methods (Figure 5C).

The final step in validation of an antibody for any application is to demonstrate that the antibody is reproducible between assay runs and between lots. We accomplish this by using a TMA containing ~120 spots that is stained with each lot of the antibody. The target expression is quantified and a regression between the two scores for each spot is performed. A high correlation among multiple lots of the antibody is our final requirement for validation. An example of this is shown for an antibody to MAP-tau (United States Biological, Swampscott, MA, USA), where two different lots of the antibody were quantified with A3UA demonstrating a high correlation (Figure 6A). Each time the antibody is used on a different TMA, sections from this control TMA are stained in parallel as an additional staining control. The control TMA slide should always have a high correlation to previous experiments for a specific and reproducible antibody. The example in Figure 6B shows inter-assay reproducibility of MENA (mouse MAb, clone 21; BD Transduction Laboratories, Franklin Lakes, NJ, USA).

These validation steps are relatively fast and inexpensive, but most importantly, they are comprehensive. This algorithm for antibody validation is somewhat similar to many other previously published approaches (13,18,36,41). These authors point out that one of the best controls for antibody validation is tissue from a knockout animal lacking the antigen altogether. When available, this tissue would be an excellent addition to any control TMA, but this is not available for every protein of interest and would be a time-consuming and expensive investment for the sole purpose of validating an antibody. Ideally, as suggested in a recent perspective by Alexander Kalyuzhny, we will one day have guidelines on the generation and use of antibodies (42) where all companies will be held to high standards for antibody validation and researchers can have greater confidence that their precious grant or clinical revenue dollars are not wasted on inaccurately labeled “vials of PBS” (36). For the meantime, however, the responsibility ultimately lies with the researcher or laboratory director to ensure that the antibodies used in their labs are validated for specificity and reproducibility.

Acknowledgments

The authors thank Elaine Alarid of the University of Wisconsin for provision of cells lines with inducible expression of ER- α . This work was supported by the National Institutes of Health (NIH; grant nos. CA139431, CA 114277, CA 110511, and CA 106709, to D.L.R.); the Susan G. Komen Foundation (grant no. KG090562, to D.L.R.); and the Department of Defense Breast Cancer Research Program [grant nos. 1W81XWH-06-1-0746 (to M.T.B.), 1W81XWH-08-1-0404 (to J.B.), and 1W81XWH-08-1-0784 (to A.W.W.)]. This paper is subject to the NIH Public Access Policy.

References

1. Dowsett M, Hanna WM, Kockx M, Penault-Llorca F, Rüschoff J, Gutjahr T, Habben K, van de Vijver MJ. Standardization of HER2 testing: results of an international proficiency-testing ring study. *Mod Pathol.* 2007; 20:584–591. [PubMed: 17396141]
2. Goldstein NS, Hewitt SM, Taylor CR, Yaziji H, Hicks DG, and Members of the Ad-Hoc Committee on Immunohistochemistry Standardization. Recommendations for improved standardization of immunohistochemistry. *Appl Immunohistochem Mol Morphol.* 2007; 15:124–133. [PubMed: 17525622]
3. Shi S-R, Liu C, Taylor CR. Standardization of immunohistochemistry for formalin-fixed, paraffin-embedded tissue sections based on the antigen-retrieval technique: from experiments to hypothesis. *J Histochem Cytochem.* 2007; 55:105–109. [PubMed: 16982846]
4. Wolff AC, Hammond MEH, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, 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.* 2007; 25:118–145. [PubMed: 17159189]

5. Hicks DG, Kulkarni S. HER2+ breast cancer: review of biologic relevance and optimal use of diagnostic tools. *Am J Clin Pathol.* 2008; 129:263–273. [PubMed: 18208807]
6. Yaziji H, Taylor CR, Goldstein NS, Dabbs DJ, Hammond EH, Hewlett B, Floyd AD, Barry TS, et al. Consensus recommendations on estrogen receptor testing in breast cancer by immunohistochemistry. *Appl Immunohistochem Mol Morphol.* 2008; 16:513–520. [PubMed: 18931614]
7. Deutsch EW, Ball CA, Berman JJ, Bova GS, Brazma A, Bumgarner RE, Campbell D, Causton HC, et al. Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE). *Nat Biotechnol.* 2008; 26:305–312. [PubMed: 18327244]
8. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumor MARKer prognostic studies (REMARK). *Breast Cancer Res Treat.* 2006; 100:229–235. [PubMed: 16932852]
9. Mullett WM, Lai EP, Yeung JM. Surface plasmon resonance-based immunoassays. *Methods.* 2000; 22:77–91. [PubMed: 11020321]
10. Franklin MC, Carey KD, Vajdos FF, Leahy DJ, de Vos AM, Sliwkowski MX. Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. *Cancer Cell.* 2004; 5:317–328. [PubMed: 15093539]
11. Hsi ED. A practical approach for evaluating new antibodies in the clinical immunohistochemistry laboratory. *Arch Pathol Lab Med.* 2001; 125:289–294. [PubMed: 11175655]
12. Leong AS. Fluorimetry in immunohistology: fact or fiction? A discussion of variables that influence results. *Appl Immunohistochem Mol Morphol.* 2004; 12:1–7. [PubMed: 15163011]
13. Michel MC, Wieland T, Tsujimoto G. How reliable are G-protein-coupled receptor antibodies? *Naunyn Schmiedebergs Arch Pharmacol.* 2009; 379:385–388. [PubMed: 19172248]
14. Jositsch G, Papadakis T, Haberberger RV, Wolff M, Wess J, Kummer W. Suitability of muscarinic acetylcholine receptor antibodies for immunohistochemistry evaluated on tissue sections of receptor gene-deficient mice. *Naunyn Schmiedebergs Arch Pharmacol.* 2009; 379:389–395. [PubMed: 18974978]
15. Jensen BC, Swigart PM, Simpson PC. Ten commercial antibodies for alpha-1-adrenergic receptor subtypes are nonspecific. *Naunyn Schmiedebergs Arch Pharmacol.* 2009; 379:409–412. [PubMed: 18989658]
16. Ramos-Vara JA. Technical aspects of immunohistochemistry. *Vet Pathol.* 2005; 42:405–426. [PubMed: 16006601]
17. Willingham MC. Conditional epitopes. is your antibody always specific? *J Histochem Cytochem.* 1999; 47:1233–1236. [PubMed: 10490451]
18. Saper CB, Sawchenko PE. Magic peptides, magic antibodies: guidelines for appropriate controls for immunohistochemistry. *J Comp Neurol.* 2003; 465:161–163. [PubMed: 12949777]
19. Pezzella F, Tse AG, Cordell JL, Pulford KA, Gatter KC, Mason DY. Expression of the bcl-2 oncogene protein is not specific for the 14;18 chromosomal translocation. *Am J Pathol.* 1990; 137:225–232. [PubMed: 2201196]
20. Willingham MC, Bhalla K. Transient mitotic phase localization of bcl-2 oncoprotein in human carcinoma cells and its possible role in prevention of apoptosis. *J Histochem Cytochem.* 1994; 42:441–450. [PubMed: 7907352]
21. Jacobs TW, Gown AM, Yaziji H, Barnes MJ, Schnitt SJ. Specificity of HercepTest in determining HER-2/neu status of breast cancers using the United States Food and Drug Administration-approved scoring system. *J Clin Oncol.* 1999; 17:1983–1987. [PubMed: 10561248]
22. Spicer SS, Spivey MA, Ito M, Schulte BA. Some ascites monoclonal antibody preparations contain contaminants that bind to selected Golgi zones or mast cells. *J Histochem Cytochem.* 1994; 42:213–221. [PubMed: 7507139]
23. Pozner-Moulis S, Cregger M, Camp RL, Rimm DL. Antibody validation by quantitative analysis of protein expression using expression of Met in breast cancer as a model. *Lab Invest.* 2007; 87:251–260. [PubMed: 17260003]
24. Grimsey NL, Goodfellow CE, Scotter EL, Dowie MJ, Glass M, Graham ES. Specific detection of CB1 receptors; cannabinoid CB1 receptor antibodies are not all created equal! *J Neurosci Methods.* 2008; 171:78–86. [PubMed: 18406468]

25. Major SM, Nishizuka S, Morita D, Rowland R, Sunshine M, Shankavaram U, Washburn F, Asin D, et al. AbMiner: a bioinformatic resource on available monoclonal antibodies and corresponding gene identifiers for genomic, proteomic, and immunologic studies. *BMC Bioinformatics*. 2006; 7:192. [PubMed: 16600027]
26. Björling E, Uhlén M. Antibodypedia, a portal for sharing antibody and antigen validation data. *Mol Cell Proteomics*. 2008; 7:2028–2037. [PubMed: 18667413]
27. Skliris GP, Rowan BG, Al-Dhaheeri M, Williams C, Troup S, Begic S, Parisien M, Watson PH, Murphy LC. Immunohistochemical validation of multiple phospho-specific epitopes for estrogen receptor alpha (ERalpha) in tissue microarrays of ERalpha positive human breast carcinomas. *Breast Cancer Res Treat*. 2009; 118:443–453. [PubMed: 19104930]
28. Sompuram SR, Kodala V, Zhang K, Ramanathan H, Radcliffe G, Falb P, Bogen SA. A novel quality control slide for quantitative immunohistochemistry testing. *J Histochem Cytochem*. 2002; 50:1425–1434. [PubMed: 12417607]
29. Saper CB. An open letter to our readers on the use of antibodies. *J Comp Neurol*. 2005; 493:477–478. [PubMed: 16304632]
30. Gustavson MD, Bourke-Martin B, Reilly D, Cregger M, Williams C, Mayotte J, Zerkowski M, Tedeschi G, et al. Standardization of HER2 immunohistochemistry in breast cancer by automated quantitative analysis. *Arch Pathol Lab Med*. 2009; 133:1413–1419. [PubMed: 19722747]
31. van der Vegt B, de Bock GH, Bart J, Zwartjes NG, Wesseling J. Validation of the 4B5 rabbit monoclonal antibody in determining Her2/neu status in breast cancer. *Mod Pathol*. 2009; 22:879–886. [PubMed: 19305385]
32. Zafrani B, Aubriot MH, Mouret E, De Crémoux P, De Rycke Y, Nicolas A, Boudou E, Vincent-Salomon A, et al. High sensitivity and specificity of immunohistochemistry for the detection of hormone receptors in breast carcinoma: comparison with biochemical determination in a prospective study of 793 cases. *Histopathology*. 2000; 37:536–545. [PubMed: 11122436]
33. Sasano H, Anderson TJ, Silverberg SG, Santen RJ, Conway M, Edwards DP, Krause A, Bhatnagar AS, et al. The validation of new aromatase monoclonal antibodies for immunohistochemistry--a correlation with biochemical activities in 46 cases of breast cancer. *J Steroid Biochem Mol Biol*. 2005; 95:35–39. [PubMed: 16024247]
34. Cheang MC, Treaba DO, Speers CH, Olivotto IA, Bajdik CD, Chia SK, Goldstein LC, Gelmon KA, et al. Immunohistochemical detection using the new rabbit monoclonal antibody SP1 of estrogen receptor in breast cancer is superior to mouse monoclonal antibody 1D5 in predicting survival. *J Clin Oncol*. 2006; 24:5637–5644. [PubMed: 17116944]
35. Brock JE, Hornick JL, Richardson AL, Dillon DA, Lester SC. A comparison of estrogen receptor SP1 and 1D5 monoclonal antibodies in routine clinical use reveals similar staining results. *Am J Clin Pathol*. 2009; 132:396–401. [PubMed: 19687316]
36. Couchman JR. Commercial antibodies: the good, bad, and really ugly. *J Histochem Cytochem*. 2009; 57:7–8. [PubMed: 18854593]
37. Mandell JW. Phosphorylation state-specific antibodies: applications in investigative and diagnostic pathology. *Am J Pathol*. 2003; 163:1687–1698. [PubMed: 14578166]
38. Mandell JW. Immunohistochemical assessment of protein phosphorylation state: the dream and the reality. *Histochem Cell Biol*. 2008; 130:465–471. [PubMed: 18648845]
39. Camp RL, Neumeister V, Rimm DL. A decade of tissue microarrays: progress in the discovery and validation of cancer biomarkers. *J Clin Oncol*. 2008; 26:5630–5637. [PubMed: 18936473]
40. Kay E, O'Grady A, Morgan JM, Wozniak S, Jasani B. Use of tissue microarray for interlaboratory validation of HER2 immunocytochemical and FISH testing. *J Clin Pathol*. 2004; 57:1140–1144. [PubMed: 15509672]
41. Rhodes KJ, Trimmer JS. Antibody-based validation of CNS ion channel drug targets. *J Gen Physiol*. 2008; 131:407–413. [PubMed: 18411328]
42. Kalyuzhny AE. The dark side of the immunohistochemical moon: industry. *J Histochem Cytochem*. 2009; 57:1099–1101. [PubMed: 19755714]

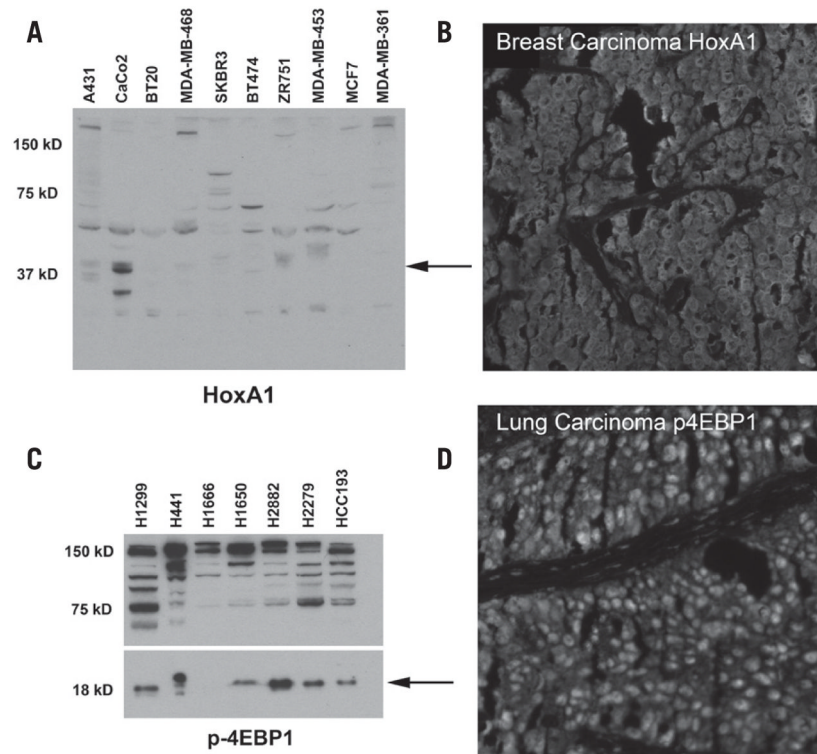


Figure 1. Examples of non-specific antibodies for HoxA1 and phospho-4EBP1

(A) Cell lysates were denatured and separated by SDS PAGE, transferred to nitrocellulose and blotted with a mouse polyclonal antibody against HoxA1. A band of the expected molecular weight is seen in CaCo2 lysate (marked by the arrow). Note the numerous bands in all cell lines at unexpected molecular weights. (B) A representative example of IHC staining on FFPE breast carcinoma tissue for HoxA1. Note the cytoplasmic staining for a nuclear transcription factor. (C) Western blot for phospho-4-EBP1. Arrow denotes band at the expected molecular weight; again, all lysates show numerous bands at unexpected molecular weights. (D) A representative example of IHC staining on FFPE lung carcinoma tissue for p-4EBP1. Note the predominantly nuclear localization for a protein expected to localize to the cytoplasm and nucleus.

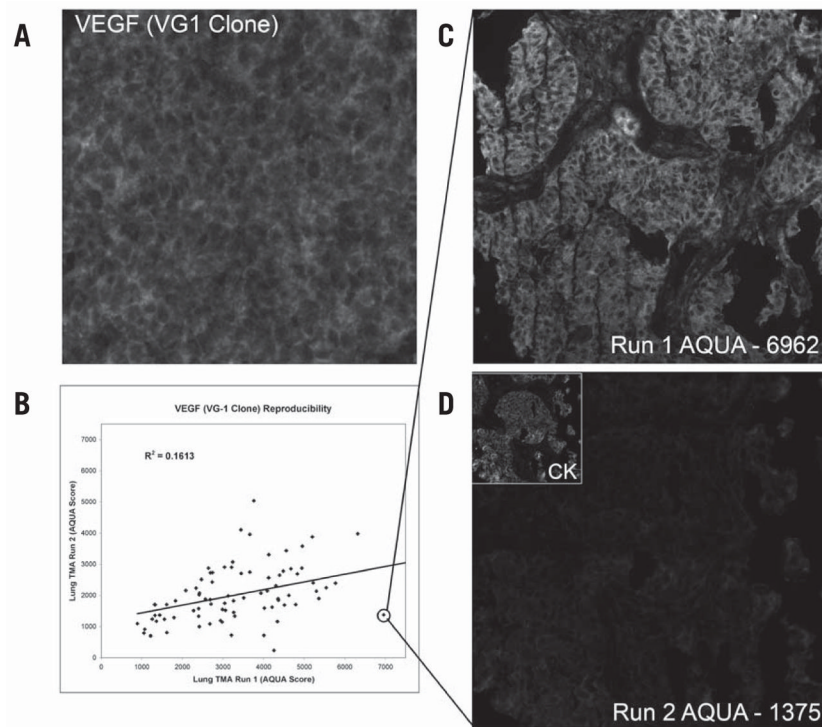


Figure 2. Lack of reproducibility is seen with the VEGF VG-1 clone antibody
 (A) IHC on FFPE lung carcinoma tissue with VG-1 (1:50 dilution) demonstrates specific staining of expected localization. (B) Serial sections of our lung TMA stained with VG-1 (1:50) do not correlate with each other. (C,D) IHC staining of serial sections of the same patient tissue showing a high level of positivity in Run 1 (C) is negative for VEGF in Run 2 (D). Inset represents cytokeratin staining of the tumor for reference.

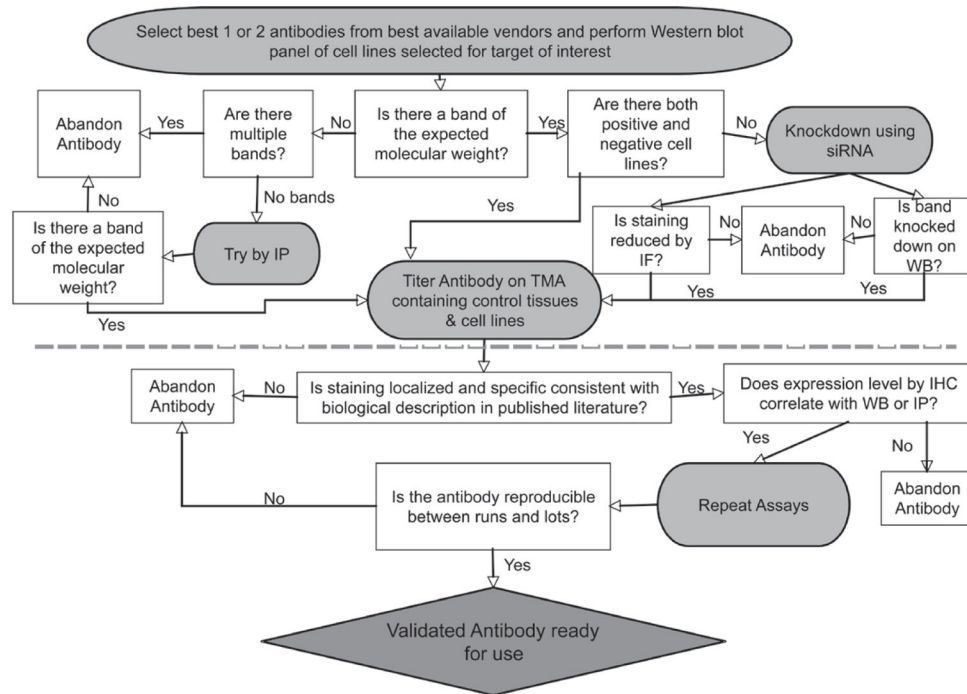


Figure 3. The Rimm Lab Algorithm for antibody validation of IHC/QIF

Step 1 of antibody validation is using cell lines in vitro to test antibody specificity. Step 2 (below the dotted line) involves further validation of antibody on tissue microarray (TMA) for expected target localization and reproducibility between assay runs and different antibody lots.

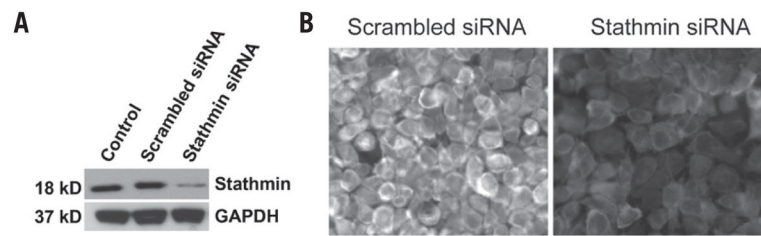


Figure 4. Stathmin expression is reduced by siRNA

(A) WB of BT-20 lysates 24 h post-mock; control scrambled siRNA or Stathmin siRNA transfection demonstrates specific reduction in Stathmin levels. GAPDH serves as a loading control. (B) BT-20 cells were fixed in 4% PFA 24 h post-transfection with scrambled control or Stathmin siRNA, and IF with Stathmin demonstrates a reduction in Stathmin expression.

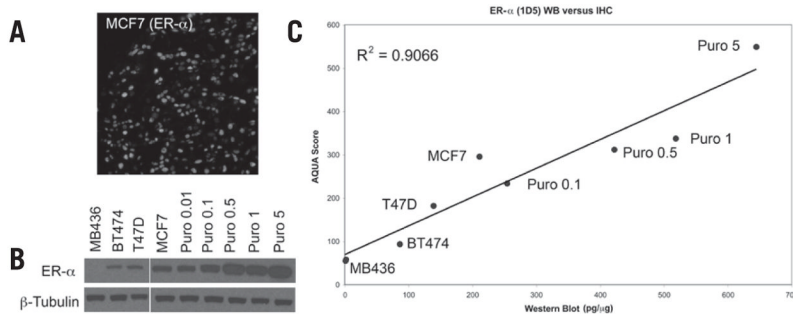


Figure 5. ER-α antibody Clone 1D5 as an example of strong correlation between quantitative WB and IHC methods

(A) FFPE MCF7 cells stained with 1D5 (1:50 dilution) demonstrate expected nuclear localization for ER-α. (B) WB with 1D5 of a panel of breast cancer cell lines demonstrating varying expression of ER-α. Puro cells are MCF7 cells stably expressing an inducible ER-α. β-tubulin serves as a loading control. (C) ER-α expression by WB was quantified using ImageJ and correlated to AQUA scores corresponding to the cell lysates demonstrating a strong correlation.

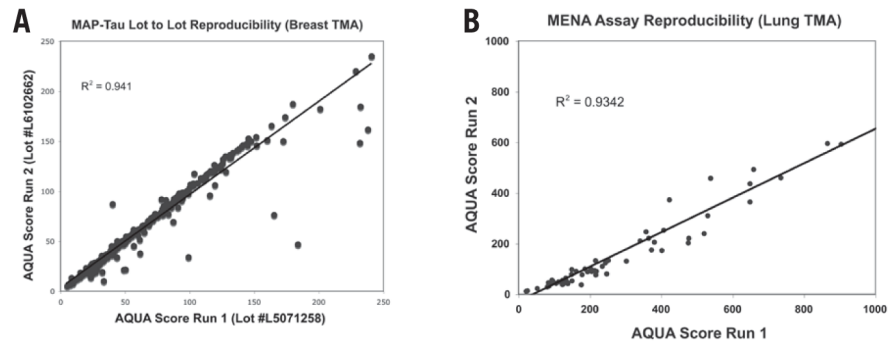


Figure 6. Examples of reproducible antibodies between assay repeats and antibody lots
(A) Lot-to-lot reproducibility is shown using a MAP-tau monoclonal antibody on a breast index TMA. (B) MENA has an R^2 value of 0.932 when stained on serial sections of our lung TMA demonstrating a high level of reproducibility among assays run on different days using the same lot and dilution of antibody.