Validation of Immunohistochemical Assays for Integral Biomarkers in the NCI-MATCH EAY131 Clinical Trial

Joseph D. Khoury1, Wei-Lien Wang2, Victor G. Prieto2, L. Jeffrey Medeiros1, Neda Kalhor2, Meera Hameed3, Russell Broaddus2, and Stanley R. Hamilton2,4

Abstract

Biomarkers that guide therapy selection are gaining unprecedented importance as targeted therapy options increase in scope and complexity. In conjunction with high-throughput molecular techniques, therapy-guiding biomarker assays based upon immunohistochemistry (IHC) have a critical role in cancer care in that they inform about the expression status of a protein target. Here, we describe the validation procedures for four clinical IHC biomarker assays—PTEN, RB, MLH1, and MSH2—for use as integral biomarkers in the nationwide NCI-Molecular Analysis for Therapy Choice (NCI-MATCH) EAY131 clinical trial. Validation procedures were developed through an iterative process based on collective experience and adaptation of broad guidelines from the FDA. The steps included primary antibody selection; assay optimization; development of assay interpretation criteria incorporating biological considerations; and expected staining patterns, including indeterminate results, orthogonal validation, and tissue validation. Following assay lockdown, patient samples and cell lines were used for analytic and clinical validation. The assays were then approved as laboratory-developed tests and used for clinical trial decisions for treatment selection. Calculations of sensitivity and specificity were undertaken using various definitions of gold-standard references, and external validation was required for the PTEN IHC assay. In conclusion, validation of IHC biomarker assays critical for guiding therapy in clinical trials is feasible using comprehensive preanalytic, analytic, and postanalytic steps. Implementation of standardized guidelines provides a useful framework for validating IHC biomarker assays that allow for reproducibility across institutions for routine clinical use. Clin Cancer Res; 24(3); 521–31. ©2017 AACR.

Introduction

Biomarker assessment is a critical component of cancer patient management. Toward that end, immunohistochemistry (IHC) plays a robust clinical role in tissue-based assessment of protein expression, particularly in solid tumors. Examples of this role include HER2 expression in breast cancer and gastroesophageal adenocarcinoma, estrogen receptor expression in breast cancer, and expression of mismatch repair (MMR) proteins in patients with colorectal adenocarcinoma or endometrial carcinoma (1–4). Nonetheless, the availability of predictive biomarkers to support therapy selection remains limited in a variety of cancer types, and this area of unmet need is expected to evolve as the use of targeted therapies, including antibodies as well as small molecules, continues to expand. As such needs expand, procedures to ensure standardized optimization and performance of such assays gain increasing importance.

Automation and technical advances in reagent chemistry, coupled with a broadening of the scope and quality of primary antibodies that recognize specific target epitopes in formalin-fixed paraffin-embedded (FFPE) tissue, have improved the quality and reliability of IHC in routine clinical practice. As with any complex in vitro diagnostic (IVD) assay, however, IHC is susceptible to variations that impact its performance, making ongoing quality monitoring a prerequisite to the clinical utility and reliability of any biomarker assessed by IHC. Nowhere is this more important than for IHC assays performed to guide therapy selection, henceforth referred to as “therapy-guiding assays.”

The Centers for Disease Control and Prevention Office of Population Genomics has provided the ACCE (analytic validity; clinical validity; clinical utility; and ethical, legal, and social implications) model (5) that can be applied to IHC (the scope of this article applies to analytic validity). The FDA categorizes IHC assays as Class II or III IVD medical devices and has provided guidance to industry (6, 7). Such devices are subject to FDA clearance or approval, albeit the latter is not a requirement for clinical use under the Clinical Laboratory Improvement Amendments of 1988 (CLIA) or the Center for Medicare and Medicare Services (CMS) requirements (8, 9). Although broad guidelines for preanalytic and analytic validation of IHC assays have been proposed (10–13) and IHC laboratory manuals are available (14), standardization and in-depth characterization of the steps required for ensuring reliable assay performance for therapy-guiding IHC biomarker assays remain...
underdeveloped. IHC assays performed in clinical trials to guide therapy are regulated by the FDA under the Investigational Device Exemptions (IDE) regulations (15).

Herein, we describe the validation principles and processes that were followed for four IHC biomarker assays [PTEN, retinoblastoma-associated protein (RB), DNA MMR protein Mlh1 (MLH1), and DNA MMR protein Msh2 (MSH2)] intended for use within the context of a nationwide clinical trial in the NCI National Clinical Trial Network, the NCI-Molecular Analysis for Therapy Choice (NCI-MATCH) EAY131 phase II precision oncology trial (NCT02465060; ClinicalTrials.gov; refs. 16, 17). Validation procedures for each of these IHC assays went beyond typical validations required for routine diagnostic IHC assays and are presented in detail as applied examples of analytic validation aimed at informing future standardization guidelines for IHC biomarker assays. All four assays described in this article were successfully validated for use in the NCI-MATCH trial.

**Primary Antibody Selection**

Primary antibodies are produced most commonly by immunization of a mammal, often mouse or rabbit, with a synthetically produced peptide portion of the target protein. Commercial products are favored over “homegrown” and customized antibodies to increase the likelihood of stable supply. Antibodies may be monoclonal or polyclonal, and the vast majority is IgG1 or IgG2 subtype. No single primary antibody attribute predetermines its performance characteristics in an IHC assay, and it is not uncommon for multiple commercially available options to be equally suitable for a given purpose.

Selection of the primary antibody starts with a thorough understanding of the target protein, its encoding gene, and the location of the epitope(s) within the broader protein structure. Ideally, the epitope is within a segment of the protein that is not (or only infrequently) impacted by mutations or deletions. In the intended application (18, 19). Selection of a primary antibody may be based further on data in published, peer-reviewed scientific studies in which sensitivity and specificity are assessed in a research setting using orthogonal protein detection methods.

Vendor attributes are important to ensure adherence to supply continuity and minimal lot-to-lot variations or interruptions. In the United States, suppliers of Class II/III IVD products are required to provide data on performance characteristics and abide by Good Manufacturing Practices (20). The antibody starting points for assay optimization are the antibody manufacturer’s recommendations and/or conditions used in peer-reviewed publications. The ideal antibody dilution may be defined as the lowest antibody concentration that yields an optimal balance between sensitivity and specificity. Commercial antigen retrieval solutions offer a pH-based choice to modulate an epitope’s three-dimensional configuration, with the aim of optimizing primary antibody binding. The ideal antigen retrieval solution minimizes exposure of tissues to stringent acidic or alkaline conditions and thus limits alteration of tissue integrity for morphologic evaluation. Incubation time should be adequate to allow antibody molecules to bind to epitope sites but short enough to minimize nonspecific binding. The use of ‘blocking’ solutions that quench endogenous peroxidase enzymes is useful for most biomarkers. The combination of staining conditions that produces an optimal signal-to-noise ratio is then adopted for the remainder of the validation process. It should be noted that FDA-approved IHC biomarker assays have preset staining parameters with minimal leeway for variations by the end user (13).

**Assay Optimization**

Next steps include determination of optimal antibody dilution, antigen retrieval conditions, and incubation time. Other parameters, such as incubation temperature, are typically preset on most automated immunostainers designed for use in the clinical environment. The use of automated platforms is strongly recommended, particularly for therapy-guiding IHC biomarker assays. The starting points for assay optimization are the antibody manufacturer’s recommendations and/or conditions used in peer-reviewed publications. The ideal antibody dilution may be defined as the lowest antibody concentration that yields an optimal balance between sensitivity and specificity. Commercial antigen retrieval solutions offer a pH-based choice to modulate an epitope’s three-dimensional configuration, with the aim of optimizing primary antibody binding. The ideal antigen retrieval solution minimizes exposure of tissues to stringent acidic or alkaline conditions and thus limits alteration of tissue integrity for morphologic evaluation. Incubation time should be adequate to allow antibody molecules to bind to epitope sites but short enough to minimize nonspecific binding. The use of ‘blocking’ solutions that quench endogenous peroxidase enzymes is useful for most biomarkers. The combination of staining conditions that produces an optimal signal-to-noise ratio is then adopted for the remainder of the validation process. It should be noted that FDA-approved IHC biomarker assays have preset staining parameters with minimal leeway for variations by the end user (13).

**Assay Interpretation Criteria**

The distribution of the target (analyte) in various tissue components determines IHC staining patterns. *A priori* knowledge of the biology of the protein provides a basis for assessing expected staining patterns (nuclear, cytoplasmic, membranous topography) and interpretation guidelines. Interpretation guidelines, including definition and handling of indeterminate results, should be defined during the validation process to ensure applicability to patient samples when an IHC assay is deployed (Table 2). Image-assisted interpretation is not under consideration for the validations illustrated in this article due to the characteristics and intended uses of the assays but is likely to have a greater role in this process over time due to improved reproducibility.

For each therapy-guiding IHC assay, a standardized reporting template is required. The template should stipulate standardized reporting terminology. For instance, the use of terms such as “positive” or “negative” can be problematic in the context of biomarkers whose loss constitutes an actionable finding. In the assessment of MMR protein expression to evaluate microsatellite instability (MSI) status, for example, negative staining for MLH1 or MSH2 is a positive result that indicates the presence of high levels of MSI (MSI-H), so the use of such terms is ambiguous.

### Table 1. Summary of primary antibodies

<table>
<thead>
<tr>
<th>Clone</th>
<th>Host</th>
<th>Isotype</th>
<th>Vendor</th>
<th>Immunogen</th>
<th>Dilution</th>
<th>Selected references</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTEN</td>
<td>Mouse</td>
<td>IgG2</td>
<td>Dako</td>
<td>Full-length protein</td>
<td>1:100</td>
<td>21</td>
</tr>
<tr>
<td>RB</td>
<td>Mouse</td>
<td>IgG2</td>
<td>EMD Millipore/Calbiochem</td>
<td>C-terminal fragment</td>
<td>1:30</td>
<td>n/a</td>
</tr>
<tr>
<td>MLH1</td>
<td>Mouse</td>
<td>IgG2</td>
<td>MilliporeSigma/Cell Marque</td>
<td>Not specified</td>
<td>1:300</td>
<td>51-53</td>
</tr>
<tr>
<td>MSH2</td>
<td>FEI1</td>
<td>IgG1</td>
<td>EMD Millipore/Calbiochem</td>
<td>C-terminal fragment</td>
<td>1:100</td>
<td>51-53</td>
</tr>
</tbody>
</table>

Abbreviation: n/a, not available.

*Peer-reviewed studies using the corresponding primary antibody.
and can lead to untoward clinical consequences. Accordingly, the terminology should be made unambiguous and describe the results of the actual assay rather than the clinical implication of the result. In the reporting templates for MLH1, MSH2, PTEN, and RB, whose loss is what constitutes an actionable finding, results are best reported as "loss of expression" and "retained expression."

**PTEN**

PTEN is a tumor-suppressor gene located on chromosome 10q23.3 encoding a dual-specificity phosphatase that acts as a dominant-negative regulator of the PI3K/AKT signaling axis. Loss of PTEN expression results in constitutive AKT activation and promotes tumor growth and altered cancer cell metabolism, mainly via upregulation of the mTOR pathway. Abnormalities in the PTEN/PI3K/AKT pathway have been detected in many human tumors, including endometrial carcinomas (21–23). PTEN knockout mice develop proliferative endometrial lesions, and germline PTEN mutations in human beings lead to Cowden syndrome (CS; ref. 24). Germline polymorphisms involving the PTEN gene have been identified in 60% to 80% of patients with PTEN hamartoma tumor-related syndromes (PHTS), a group of disorders that includes Cowden syndrome, Bannayan–Riley–Ruvalcaba syndrome (BRRS), and PTEN-related Proteus syndrome and Proteus-like syndrome. All of these syndromes are characterized by tissue overgrowth and benign tumors, but only Cowden syndrome and BRRS have a predisposition for cancer development.

In tumors with PTEN loss, the activated PI3K/AKT/mTOR pathway constitutes an attractive target of therapy. Loss of PTEN expression detected by IHC is regarded as the most accurate reflection of the loss of PTEN function and, as a result, serves to determine eligibility for therapies that target critical downstream nodes within the PI3K/AKT/mTOR axis in the NCI-MATCH trial.

**RB**

The RB transcriptional corepressor 1 (RB1) gene is a tumor suppressor gene that encodes the RB protein, a negative cell-cycle regulator (27). Mutations and deletions of RB1 are common in many cancers, and inherited allelic loss of RB1 confers increased cancer susceptibility. The RB protein and its two family members, p107 and p130, regulate cell proliferation through transcriptional repression of genes involved in cell-cycle transition from the G1–S phase (27). Loss of RB function allows unregulated cell-cycle progression and promotes tumor growth. Specifically, cell-cycle progression requires the dissociation of the RB/E2F complex, which is tightly regulated physiologically via RB phosphorylation. In tumors, constitutive

---

**Table 2. Interpretation guidelines**

<table>
<thead>
<tr>
<th>Staining pattern</th>
<th>Positive therapy-guiding alteration</th>
<th>Lack of therapy-guiding alteration</th>
<th>Indeterminate results</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTEN Cytoplasmic</td>
<td>Loss of expression (absent)</td>
<td>Retained expression (retained)</td>
<td>• Nonneoplastic elements weak* • Incongruent staining pattern • Insufficient sample</td>
</tr>
<tr>
<td>RB Nuclear</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLH1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSH2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Staining should be repeated in such a situation before indeterminate category is assigned.
disruption of the RB/E2F complex results from loss of RB expression through deletions or mutations, or from increased RB phosphorylation. With abrogation of RB-dependent cell-cycle inhibition, cell-cycle transition and commitment to cell division are coordinated by cyclin-dependent protein kinases (CDK), which are emerging targets for therapy (28). Palbociclib is a potent selective inhibitor of CDK4 and CDK6, with significant activity in breast cancer models (29, 30). Of note, palbociclib shows no activity in RB-deficient cells (28).

RB is a nuclear protein that is expressed ubiquitously in human tissues. Neoplasms with retained RB expression show nuclear expression by IHC, whereas those with RB loss have neoplastic cells that distinctly lack nuclear reactivity (Fig. 2). Aside from the nuclear staining pattern, interpretation guidelines are similar to those detailed above for PTEN.

**MLH1 and MSH2**

The genes mutL homolog 1 (MLH1) and mutS homolog 2 (MSH2) are located on chromosome 3p21.3 and chromosome 2p21, respectively. Both belong to a family of genes known as MMR genes. MLH1 encodes a protein which heterodimerizes predominantly with PMS2, a mutL homolog, to form the mutL complex. MSH2 encodes a protein that heterodimerizes predominantly with MSH6, a mutS homolog, to form the mutS complex. Both complexes are essential for the detection and initiation of repair of DNA strand misalignment and base-pair matching errors that occur during DNA replication. Loss of function of either MLH1 or its binding partner PMS2 or of MSH2 and its binding partner MSH6 results in error-prone DNA replication. This abnormality leads to, among other effects, alterations in the length of tandem DNA sequence repeats called microsatellites, a condition known as MSI or deficient MMR (dMMR). Uncorrected mutations occur throughout the genome, termed “hypermutation” or “tumor mutation burden.” The extent of microsatellite alterations is assessed semiquantitatively as MSI-low or MSI-high in DNA-based assays, with the latter having a higher degree of correlation with bona fide MMR genotypic defects.

In most (~95%) cases, loss of MMR results from sporadic methylation–induced MLH1 promoter inactivation of both copies of the gene that can occur in a wide variety of tumors, especially colonic and endometrial adenocarcinomas (28). Somatic mutations occur occasionally. Germline mutations in MMR genes, most of which involve MLH1 or MSH2, result in the autosomal-dominant Lynch syndrome, formerly hereditary nonpolyposis colorectal cancer syndrome (HNPPC). These patients have a significantly increased risk for colorectal, gastric, small intestinal, liver, gallbladder, urothelial, brain, and skin tumors. Women also have an increased risk of ovarian and uterine endometrial carcinomas. Muir–Torre syndrome is a subtype of Lynch syndrome in which patients have skin neoplasms (sebaceous tumors and keratoacanthomas) in addition to an increased risk of developing visceral malignancies. Tumors with MLH1 or MSH2 loss, or other MMR defects, have distinctive clinical features compared with those without MMR defects. For example, colonic adenocarcinomas with MMR defects (e.g., loss of MLH1, MSH2, etc.) have a better stage-specific prognosis, and these patients do not benefit from 5-fluorouracil monotherapy (31, 32).

In view of the prognostic, therapeutic, and genetic implications of MMR aberrations in cancer, guidelines have been in place for many years to screen for and delineate their nature. In accordance with the commonly used guidelines from several professional organizations, screening is typically carried out by surveying the expression of MLH1, PMS2, MSH2, and MSH6 in tumor tissue using IHC. Tumors that exhibit loss of expression of one or more MMR genes are often reflexed for MSI evaluation using molecular techniques. Genetic counseling and genomic evaluation then may be pursued in line with the findings in the tumor and family history to identify family members with Lynch syndrome (4).

Tumors with retained MLH1 or MSH2 expression and nonneoplastic cells have diffuse moderate-to-strong nuclear staining by IHC, whereas tumors that have lost MLH1 or MSH2 expression have a distinctive absence of nuclear expression (Fig. 3). These tests were the best way to determine if patients had MMR, which had been shown to respond to PD-1 checkpoint inhibitors, a treatment option in NCI-MATCH (33, 34). Occasional cases have a speckled nuclear pattern on IHC for MLH1, with complete loss of expression of the binding partner PMS2 and the presence of MSI-H/dMMR. Aside from this rare staining pattern, interpretation guidelines are similar to those detailed above for PTEN.

**Indeterminate results**

In the interpretation of biomarker results, occasional cases will have difficult-to-interpret and/or uncertain results due to technical or nontechnical variances that may be attributable to

---

**Figure 2.**

Examples of positive and negative RB expression by IHC. **A** and **B** (100×). Positive RB expression. The nuclei of nonneoplastic colonic mucosa and submucosa (**A**) and colonic adenocarcinoma (**B**, white arrow) exhibit diffuse strong nuclear labeling. **C** (100×). Loss of RB expression in lung adenocarcinoma.

Loss of nuclear staining in neoplastic cells (blue counterstain nuclei, black arrow), whereas adjacent nonneoplastic stromal cells and lymphocytes have retained nuclear expression (dark brown nuclei, white arrow).
preanalytic, analytic, or postanalytic factors. Thus, validation documents and subsequent standard operating procedures need to address the processes by which indeterminate/equivocal results are handled. Absent or weak labeling of non-neoplastic elements (internal controls) usually indicates a preanalytic issue (e.g., tissue fixation or processing, or inadequate sampling by biopsy) or an analytic problem in the IHC staining process. In such instances, the test may be repeated to exclude a possible one-off issue. If unresolved, the sample may be regarded as indeterminate. The indeterminate category should also include situations in which tissue is insufficient, for example, due to extensive necrosis or inadequate sampling.

Criteria for handling cases with heterogeneous staining results also should be clarified. For example, in cases with heterogeneous staining for PTEN or RB, wherein only a subset of neoplastic cells shows loss of expression, the antigen may be regarded as retained as the cellular subsets with retention would not be responsive, respectively, to AKT- or CDK4/6-targeted therapies.

Validation Samples

Cell lines

Cell lines with known mutation and expression profiles can offer a useful tool for orthogonal validation after the IHC assays are finalized and locked for use. Preparation of cell pellets that are fixed in 10% neutral-buffered formalin and embedded in paraffin can be used to perform IHC assays, analogous to cytology cell blocks. As cell lines are a convenient source of well-characterized cells that can be used for a variety of DNA and RNA assays, the performance of IHC on FFPE cell line blocks allows direct comparison of an IHC assay to other protein-level or DNA/RNA-level assays. A list of cell lines used in our validations is in Table 3.

The PTEN-mutant (mut) and wild-type (wt) cell lines were obtained from The University of Texas MD Anderson Cancer Center (MDACC) Characterized Cancer Cell Line Core Facility. IHC results for the PTEN mut cell line (MDA-MB-468) demonstrated absence of protein staining, whereas the PTEN wt cell line

<table>
<thead>
<tr>
<th>Table 3. Summary of cell lines</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cell line</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>PTEN</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>RB</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>MLH1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>MSH2</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Validation of IHC Integral Marker Assays

Figure 3.
Examples of positive and negative cases for MLH1 expression by IHC. A and B (100 x), Positive MLH1 expression. The nuclei of colonic adenocarcinoma (A) and endometrial endometriod adenocarcinoma (B) exhibit diffuse strong staining. C and D (100 x), Loss of MLH1 expression in colonic adenocarcinoma (C) and endometrial endometriod adenocarcinoma (D). Loss of nuclear staining in neoplastic cells (black arrow), whereas adjacent normal stromal cells and lymphocytes have retained nuclear expression (white arrow). Similar staining patterns are seen for MSH2 (not shown).
(MDA-MB-231) demonstrated retained PTEN expression. These results were identical to those previously obtained using reverse protein phase array expression studies (data not shown).

Similarly, the RB1<sup>mut</sup> cell line A2058 (derived from melanoma) demonstrated complete absence of RB staining, whereas the cell lines A549 (lung carcinoma) and BJ (fibroblast) with RB1<sup>wt</sup> demonstrated positive staining (Fig. 4). These cell lines were purchased from the ATCC.

The MLH1<sup>mut</sup> cell line HCT116 demonstrated complete absence of MLH1 staining, whereas the cell line SW480 with MLH1<sup>wt</sup> had positive staining by IHC. Similarly, the MSH2<sup>mut</sup> cell line LoVo demonstrated complete absence of MSH2 staining, and the SW480 with MSH2<sup>wt</sup> had retained staining. These cell lines were purchased from the ATCC.

Patient samples

A major component of clinical validation of IHC assays entails the use of patient samples as the source of materials to demonstrate an appropriate spectrum of staining across a particular tumor type or diverse tumor types. The validations we performed included a main cohort of tumors of specific histologic types commonly evaluated with the particular assay (e.g., colorectal adenocarcinoma for MLH1/MSH2) as well as a wider array of available tumor samples with molecular and/or FISH results that were used for reasons of convenience. In addition, when feasible, samples from selected patient cohorts (e.g., Cowden syndrome, Muir–Torre syndrome, etc.), wherein the performance of the assay is predictable, were also used to provide further confirmation of assay performance and reliability.

Validation Results

Accuracy

Sensitivity and specificity. The diagnostic accuracy of a laboratory assay is defined as the extent of its agreement with a reference standard, with a reference standard being "the best available method to establishing the presence or absence of the target condition" (35–37). Selection of a reference standard for a particular assay is predicated on the biology of the analyte being measured and the inherent characteristics of the tool used to measure it. Estimates of sensitivity and specificity are among the most commonly employed measures of diagnostic accuracy. On the basis of definitions adopted by the FDA (35) from the Clinical Laboratory Standards Institute Harmonized Terminology Database (36) and the Standards for Reporting of Diagnostic Accuracy (STARD) initiative (37), sensitivity of a test is the "proportion of subjects with the target condition in whom the test is positive," whereas specificity is the "proportion of subjects without the target condition in whom the test is negative." (35).

Cross-validation of PTEN. Analytic performance of the PTEN IHC assay was determined through cross-validation by measuring positive and negative concordances with an IHC PTEN assay.
previously validated at another institution, the Memorial Sloan Kettering Cancer Center (MSKCC). As such, the MSKCC assay was considered the benchmark (comparator) against which the performance of the MDACC PTEN IHC assay was compared for the purposes of analytic validation. The selection of this benchmark was based on the fact that genomomic alterations that lead to loss of PTEN expression include biallelic loss of or mutations in the PTEN gene whose detection requires FISH and next-generation sequencing (NGS), respectively, neither of which has been shown to be superior to IHC in identifying loss of PTEN function (25, 26). The PTEN IHC assay at MSKCC is performed on identical autostainers (Leica Biosystems) using the same antibody clone and titer (Dako; clone 6H2.1; 1:100). However, notable differences between the MSKCC and MDACC staining protocols included the duration of antigen retrieval time (30 minutes vs. 20 minutes, respectively) and primary antibody incubation time (30 minutes vs. 15 minutes, respectively), which produced a slightly different signal intensity. To validate the MDACC PTEN IHC assay, 31 sections from FFPE blocks containing tumor tissue were stained at MSKCC and interpreted by pathologists at both institutions without knowledge of the results at the other institution. Results were concordant in 29 of 31 (94%) cases [95% confidence interval (CI), 0.771–0.989], with an analytic sensitivity of 92% (95% CI lower limit, 0.715) and specificity of 100% (95% CI lower limit, 0.560; Supplementary Table S1). Analysis of discordant results in two false-negative cases relative to comparator was included in the validation report. These two cases had moderate to weak labeling on the MSKCC IHC, which was either not noted on MDACC stains or attributed to nonspecific staining at the tissue edge or within areas of necrosis. One of the two discordant cases assessed for mutations using a clinically validated NGS mutation screening panel (38) was found to be negative for PTEN mutations.

Cross-validation of RB. For the RB IHC assay, RB1 mutation analysis using the aforementioned clinically validated NGS panel was considered the benchmark comparator. Forty-eight human tumors and six normal human control tissues were selected from the pathology files at MDACC on the basis of available mutation results and adequate residual tissue specimens. The selection of this benchmark was based on the fact that cases with RB1 mutations were expected to lack RB protein expression (29, 39). Using these criteria, the RB IHC biomarker assay had a sensitivity and a specificity of 100% (95% CI lower limit, 0.828 and 0.858, respectively), which produced a slightly different signal intensity. To validate the MDACC PTEN IHC assay, 31 sections from FFPE blocks containing tumor tissue were stained at MSKCC and interpreted by pathologists at both institutions without knowledge of the results at the other institution. Results were concordant in 29 of 31 (94%) cases [95% confidence interval (CI), 0.771–0.989], with an analytic sensitivity of 92% (95% CI lower limit, 0.715) and specificity of 100% (95% CI lower limit, 0.560; Supplementary Table S1). Analysis of discordant results in two false-negative cases relative to comparator was included in the validation report. These two cases had moderate to weak labeling on the MSKCC IHC, which was either not noted on MDACC stains or attributed to nonspecific staining at the tissue edge or within areas of necrosis. One of the two discordant cases assessed for mutations using a clinically validated NGS mutation screening panel (38) was found to be negative for PTEN mutations.

Cross-validation of MLH1 and MSH2. For the MLH1 IHC assay, MSI testing and MLH1 gene promoter methylation were considered the benchmark comparators against which MLH1 expression by IHC was assessed. For the MSH2 IHC assay, MSH2 testing was considered as the benchmark comparator. The selection of these benchmarks for MLH1 and MSH2 was based on guidelines and recommendations based on specific patterns of inactivation of the genes encoding each of these MMR proteins (40, 41). Two cases had no MSI data, but patients were verified to harbor MSH2 germline mutations through germline sequencing after genetic counseling. It was our premise that a tumor that is microsatellite stable would not be expected to have a deleterious mutation in an MMR gene, including MLH1 and MSH2. Forty tumors comprised of known MLH1 (20 cases) or MSH2 (20 cases) loss were selected from the surgical pathology archives of MDACC for the validation studies. The validation set consisted of whole-tissue sections of various tumor types, including cases of colorectal, endometrioid, and esophageal adenocarcinoma; urothelial carcinoma; and tubulovillous adenomas.

Sensitivity was defined as the percentage of samples that had loss of MLH1 or MSH2 expression and classified as MSI-high, and that had respectively MLH1 promoter methylation or MSH2 mutation. Specificity was defined as the percentage of samples that had retained MLH1 or MSH2 expression and that were classified as MSI-stable. The sensitivity and specificity of the IHC MLH1 and MSH2 assays were 100% (95% CI lower limit, 0.799 for each; Supplementary Tables S4–S6).

Selected samples from specific relevant clinical context. Samples from patients with a specific clinical context (CS, retinoblastoma, hereditary osteosarcoma, hereditary osteosarcoma) were identified from the pathology files at MDACC. Electronic medical records were reviewed for pathologic and clinical features. FFPE tissue blocks were sectioned, and slides were stained per respective IHC protocols.

PTEN IHC assay performance in tumors from patients with CS. CS is an autosomal-dominant genetic syndrome characterized by germline mutations in PTEN manifested by multiple benign skin tumors (including sclerotic fibromas and trichilemmomas), gastrointestinal hamartomatous polyps, and an increased risk for various cancers (including those arising in the thyroid, endometrium, breast, and kidney). Earlier studies suggested that tumors associated with CS have loss of PTEN expression assessed by IHC (42).

Specimens from three patients tested previously to have PTEN germline mutations were determined to have CS by MDACC genetic counselors. Tumors tested include metastatic breast carcinoma, trichilemmoma, and sclerotic fibroma. As expected, all tumors (3/3, 100%) had complete loss of PTEN by IHC, with positive internal nonneoplastic tissue controls.

RB IHC assay performance in retinoblastomas and retinoblastoma-associated osteosarcoma. Retinoblastoma is rare cancer that arises from the retina and most often occurs in early childhood. Retinoblastoma may be hereditary (40%) or sporadic (60%), with hereditary tumors developing in infants and often bilaterally. These tumors characteristically harbor deleterious biallelic (or homozygous) mutations in RB1 (most common) or have deletions of both RB1 loci located on chromosome 13q14. Patients with germline RB1 mutations are susceptible for the development of other tumors, including osteosarcoma and urothelial carcinoma.

Specimens from seven patients included retinoblastomas from six patients and a right tibia osteosarcoma from a patient with a history of early-onset bilateral retinoblastomas. As expected, all tumors (7/7, 100%) had complete loss of RB expression by IHC, with positive internal nonneoplastic tissue controls.

MLH1 and MSH2 IHC assay performance in tumors from patients with Lynch and Muir–Torre syndromes. Muir–Torre syndrome is an autosomal-dominant genetic variant of Lynch syndrome characterized by germline mutations in MMR genes, including MLH1 and MSH2. These patients have multiple benign skin tumors (sebaceous tumors and keratoacanthomas) and an
increased risk for visceral cancers. Prior studies suggest that
tumors in patients with Lynch and Muir–Torre syndromes
with germline mutations in MLH1 or MSH2 have loss of MLH1
or MSH2 expression, respectively (43–46).

Four patients with germline MLH1 (n = 2) or MSH2 (n = 2)
mutations were selected by MDACC genetic counselors. Seba-
ceous adenomas from three patients and an urothelial carcinoma
from one patient were tested. As expected, all tumors (4/4, 100%)
had complete loss of MLH1 or MSH2 as assessed by IHC, with
positive internal nonneoplastic tissue controls.

Precision
Interpathologist scoring concordance

MLH1 and MSH2 IHC assays performed on the main patient set
(n = 40) were interpreted independently by two pathologists (J.D.
Khoury and W.-L. Wang) according to interpretation guidelines
summarized in Table 2. Results were concordant (40/40, 100%;
95% CI, 0.943–1.000) for both assays, and the interpretation
guidelines were applicable to all cases evaluated.

External proficiency testing
As part of ongoing quality assurance and improvement, the
MDACC clinical IHC laboratory participates in a nationwide,
external quality assessment survey, the DNA MMR Proficiency
Testing (2011–current). The CAP administers this proficiency test
twice a year. Unstained FFPE slides from a single sample are
provided to laboratories to perform MMR IHC studies, including
MLH1 and MSH2. Laboratories are blinded to the MLH1 and
MSH2 expression status and must render interpretations based on
their in-house assays. The results are then compared with other
participating and reference laboratories. In all 5 years, the results
of MLH1 and MSH2 staining performed at MDACC were con-
cordant with testing performed at all participating institutions,
with no unacceptable interpretations rendered. No commercial
proficiency tests exist for PTEN and RB IHC assays.

Discussion

In this article, we provide details of validation procedures for
four IHC assays currently used in conjunction with molecular
tests to guide therapy for patients enrolled in the NCI-MATCH trial (Fig. 5). The aim of these procedures was to establish the clinical validity of the biomarker tests in question as adjunct tools for targeted therapy and to ensure that their performance characteristics meet acceptable quality limits and standards. In addressing the various facets of therapy-guiding IHC biomarker assay development in a way that exceeds typical validation procedures for clinical diagnostic assays, the approach outlined in this article strives to ensure that no details are omitted in the deployment of such assays in the clinical setting.

We have proposed the creation of a specialized accreditation program to create certified advanced companion diagnostics facilities (47–49). The integration of stringently validated therapy-guiding IHC biomarker assays into such a paradigm can provide further refinement of personalized therapy selection in specific contexts, particularly where genomic or epigenetic lesions that cause loss of expression of a particular biomarker (e.g., MLH1 or PTEN) might not be readily detectable by mutation screening assays (Fig. 6). A similar recommendation to strengthen the oversight and accreditation of laboratories performing biomarker tests for targeted therapies was recently advocated by the Health and Medicine Division (formerly, Institute of Medicine), a division of the National Academies of Sciences, Engineering, and Medicine (50).

No specific guidelines exist for standardized validation of therapy-guiding biomarker assays. Components of the validation of the four IHC assays described here were collated from various sources and represent an iterative process that was developed in collaboration with an advisory team overseeing biomarker support for the NCI-MATCH trial. It is hoped that the validation steps outlined herein may provide a blueprint to inform development of standardized validation guidelines that will gain wide acceptance by regulatory agencies and payers. Ideally, the blueprint would be developed through collaborative efforts between key stakeholders, including the pathology and molecular diagnostics communities, oncologists, and the pharmaceutical industry.

Figure 6.
Screening-refinement model for the selection of targeted therapies. Next-generation companion diagnostics, which include NGS and possibly large-scale proteomics, provide a screening platform for potential therapy selection. Precision therapy-guiding companion diagnostics, mostly including IHC, may provide further refinement of inclusion/exclusion decisions for targeted therapies.
**Disclosure of Potential Conflicts of Interest**

I.D. Khoury reports receiving commercial research grants from Angle, Kinexus, and Sonotent. S.B. Hamilton is a consultant/advisory board member for Halio DX, LOXO Oncology, and Merck. No potential conflicts of interest were disclosed by the other authors.

**Acknowledgments**

The authors thank the staff of the Clinical Immunohistochemistry Laboratory at The University of Texas MD Anderson Cancer Center, especially Janet Quiñones and Victor Ortega; the NCI-MATCH support team at MD Anderson, especially Dr. Geeta Mantha, Dr. Mark Roubort, Dr. Ignacio Wistuba, Dr. Jaime Rodriguez-Canales, and Barbara Mino for assistance with external validation studies for PTEN. The authors also thank the NCI-MATCH Manuscript Committee for editorial review of this manuscript. This study was supported by the NIH/National Cancer Institute award U10CA180820 and Cancer Center Support Grant award P30 CA016672.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 4, 2017; revised July 6, 2017; accepted August 9, 2017; published OnlineFirst August 24, 2017.

**References**


Validation of Immunohistochemical Assays for Integral Biomarkers in the NCI-MATCH EAY131 Clinical Trial


Access to the most recent version of this article at:
doi:10.1158/1078-0432.CCR-17-1597

Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2017/08/24/1078-0432.CCR-17-1597.DC1

This article cites 44 articles, 10 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/24/3/521.full#ref-list-1

This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/24/3/521.full#related-urls

Sign up to receive free email-alerts related to this article or journal.

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

To request permission to re-use all or part of this article, use this link
http://clincancerres.aacrjournals.org/content/24/3/521.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.