Introduction

This presentation will discuss the assessment of prognostic biomarkers by immunohistochemistry (IHC), using HER2 and estrogen receptor alpha (ER) in breast cancer as examples to illustrate important issues relating to utility, validation, problems, and solutions. In general, assessing almost any prognostic biomarkers by IHC is more challenging than assessing diagnostic biomarkers because life-saving treatment decisions in every patient rely on highly accurate, reproducible, and quantitative results, which enzymatic IHC is not particularly well suited for [1]. Nonetheless, nearly all testing for HER2 and ER in breast cancer is currently based on enzymatic IHC, and it is our responsibility as pathologists to ensure that they meet the highest standards possible.

HER2 Testing by IHC

HER2 (also referred to as HER2/neu and c-erbB2) is a proto-oncogene located on chromosome 17. It encodes a tyrosine-kinase protein residing on the outer cell membrane. Many studies during the past 20 years have shown that the HER2 gene is amplified in 15-30% of IBCs (closer to 15% today, which is probably due to screening mammography), and that amplification is highly correlated with over-expression of the protein [2-4]. The relationship between HER2 status and clinical outcome is complex. There is a weak but significant association between poor outcome and “positive” (i.e. amplified and/or over-expressed) HER2 in patients receiving no additional therapy following initial surgery [5, 6]. Most patients receive some type of adjuvant therapy and the association between HER2 status and outcome in this setting appears to depend on the type of therapy. Currently, the most important are antibody-based therapies targeting the HER2 protein, such as trastuzumab, and many studies (including randomized clinical trials) show that HER2 positive tumors are responsive to trastuzumab in both adjuvant and metastatic settings [7, 8]. Because of these beneficial clinical results, HER2 testing is now mandatory in all patients with breast cancer [9, 10], and IHC (which measures protein over-expression) and fluorescent in situ hybridization (FISH; which measures gene amplification) are the two primary methods being utilized, especially IHC. International proficiency testing programs and some of the same studies which validated the clinical efficacy of trastuzumab noted that the error rate in testing with IHC was as high as 20% (including both false-negative and false-positive results) [11-13]. The error rate associated with FISH was less clear, but there was an impression held by many that it was more reliable than IHC. It soon became fairly clear that this unacceptably high error in HER2 testing by IHC was widespread, which motivated many professional medical societies to develop strategies to improve it. Among the most notable outcome of these efforts to date was the publication of testing guidelines in late 2007 which were developed jointly by the College of American Pathologists (CAP) and American Society of Clinical Oncology (ASCO) [14, 15]. These guidelines, which addressed both IHC FISH, were based on a comprehensive review of the medical literature by panels of experts, and have resulted in many changes for pathologists performing these tests (see Figure 1). For example, laboratories performing HER2
testing are required to strictly follow the guidelines and perform comprehensive ongoing quality assurance programs to obtain CAP accreditation. Hopefully, due to these and other efforts, we will see a dramatic improvement in the quality of HER2 testing in the near future. It is very clear that meeting the guidelines and obtaining accurate results requires substantial training and experience, and that laboratories lacking either should not be performing the tests at all.

![Figure 1. Algorithms for HER2 testing in breast cancer by immunohistochemistry (left) and fluorescence in situ hybridization (right) [14, 15].](image)

**ER Testing by IHC**

ER is a nuclear transcription factor activated by estrogen to regulate growth and differentiation of normal breast epithelial cells [16-18]. These pathways remain operative to varying degrees in IBCs, including estrogen-stimulated growth of tumor cells expressing ER, which is detrimental [17-19]. ER expression has been measured in IBCs for almost 40 years. During the first 20-25 years it was measured by radio-labeled biochemical ligand (i.e., estrogen)-binding assays (LBAs) on whole tissue extracts prepared from fresh-frozen tumor samples, which was costly and difficult. Many studies using LBAs in large randomized clinical trials demonstrated that ER was a relatively weak prognostic factor but a very strong predictive factor for response to hormonal therapies such as tamoxifen [19], which is one of the most widely used types of hormonal therapy. Tamoxifen, which binds ER and blocks estrogen-stimulated growth, was shown to significantly reduce disease recurrence and prolong life in patients with ER-positive IBCs [19, 20], which is also true for newer types of hormonal therapies such as aromatase inhibitors [21, 22]. The primary reason for assessing ER is its ability to predict response to hormonal therapies.

Although the clinical utility of assessing ER was initially based almost entirely on studies using technically standardized LBAs, in the mid-1990s laboratories around the world began abandoning LBAs in favor of IHC, and IHC is used for nearly all testing today, and the validation of this strategy is problematic and evolving. Notwithstanding, there are several advantages associated with IHC compared to LBAs, especially its ability to measure ER on routine formalin-fixed paraffin-embedded samples, eliminating the need for fresh-frozen samples and the onerous infrastructure required to provide it. Other advantages include lower cost, better safety, and superior sensitivity and specificity in the sense that the assessment of ER is restricted to tumor cells under direct microscopic visualization, independent of tumor...
Immunohistochemistry in Routine Clinical Testing of Prognostic Biomarkers

Utility, Validation, Problems, and Solutions

D. Craig Allred, MD, FCAP

cellularity or the presence of benign epithelium, which could be problematic for LBAs. For all these reasons and more, IHC was approved by the CAP and ASCO for routine clinical use [9, 10]. Despite these approvals, however, there are significant problems with IHC that persist today, including the widespread use of diverse staining procedures of unequal quality and varied often arbitrary methods of interpreting results, resulting in an overall error rate estimated to be as high as 20% [23-26] - which has created a "crisis" in ER (and PgR) testing in the minds of many [27]. In response, the CAP and ASCO recently convened another panel of experts to review the literature and develop guidelines for ER (and PgR) testing by IHC. Although the results are still forthcoming, it is likely that they will be similar to the HER2 guidelines. In general, all good guidelines agree that tests used in routine clinical practice should be based on sensitive and specific reagents, standardized laboratory procedures and, especially, calibrated to relevant clinical outcome in a comprehensive manner [9, 28, 29].

References


Allred - 3 -


Quality Assurance, Error Reduction, and Patient Safety in Anatomic Pathology

Immunohistochemistry in Routine Clinical Testing of Prognostic Biomarkers

Utility, Validation, Problems, and Solutions

Examples:
HER2 and ERα in Breast Cancer

D. Craig Allred, MD, FCAP
Department of Pathology and Immunology

Washington University in St. Louis
School of Medicine

© 2009 College of American Pathologists. Materials are used with the permission of the faculty.
Outline

Define validation of routine prognostic factors.

Example: HER2 testing by IHC

- Overview/Utility
- Validation
- Problems
- Solutions (CAP/ASCO Guidelines 2007)

Example: ERα testing by IHC

- Overview/Utility
- Validation
- Problems
- Solutions (impending CAP/ASCO Guidelines 2009)

General principles and practices of “good” testing
General Guidelines for Evaluating Prognostic Factors in Routine Clinical Practice

Clinical Validation: The factor(s) should identify groups of patients with significantly different risks of relapse, survival, or treatment response - ideally demonstrated in multiple randomized studies.

Technical Validation: The assays should be specific, sensitive, reproducible, calibrated to clinical outcome, interpreted, and reported in a relatively uniform manner. There should be comprehensive ongoing quality assurance.

Useful: Actually used by physicians to make important treatment decisions.

Overview
Testing for HER2 in Breast Cancer

Many Assays
- W.Blot
- IHC
- ELISA
- S.Blot
- FISH
- CGH
- CISH
- N.Blot
- RT-PCR
- Others...

Many Assays/ABs
- HercepTest (PAb)
- CB11
- Tab250
- PAB1
- 21N
- SP3
- Others...

Protein Expression
Gene Amplification
RNA Expression

Protein
Expression

Gene
Amplification

RNA
Expression

Highly Correlated

Amp=↑ Exp

Detection Systems
Non-Linear Amplification

Many Purposes
- Prognostic
  - Untreated (+ = bad)
- Predictive
  - CMF (+ = bad)
  - Anthracylines (+ = good)
  - Endocrine (+ = bad)
  - Taxanes (+ = good)
  - Herceptin (+ = good)
HercepTest IHC Scoring Method

<table>
<thead>
<tr>
<th>Staining Pattern</th>
<th>Score</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>No staining.</td>
<td>0</td>
<td>Negative</td>
</tr>
<tr>
<td>Faint incomplete staining of cell membrane in &gt;10% of tumor cells.</td>
<td>1+</td>
<td>Trace Negative</td>
</tr>
<tr>
<td>Weak to moderate complete staining of cell membrane in &gt;10% of tumor cells.</td>
<td>2+</td>
<td>Weak Positive</td>
</tr>
<tr>
<td>Strong complete staining of cell membrane in &gt;10% of tumor cells.</td>
<td>3+</td>
<td>Strong Positive</td>
</tr>
</tbody>
</table>

*Arbitrary strategy which does not convey the true (substantial) heterogeneity of expression, and prone to false-positive results due to many technical artifacts.*
Expected “True” Distribution of IHC (HercepTest) Results

<table>
<thead>
<tr>
<th></th>
<th>0/1+</th>
<th>2+</th>
<th>3+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mayo Clinic (n=1,556)</td>
<td>73%</td>
<td>14%</td>
<td>13%</td>
</tr>
</tbody>
</table>

A Potential Problem with IHC (HercepTest): False-Negatives

Central Lab IHC =3+
Local Lab IHC <3+ (n = 268): 15%

HER First Trial; First-line Herceptin + other in metastatic/HER2+ pts
See: Breast Cancer Res Treat 76:S68(abst#235), 2002

*Consequence = Under Treatment*
A Potential Problem with IHC (HercepTest): False-Positives

| Central Lab IHC <3+ |  
|---------------------|---|
| ¹Local Lab IHC =3+ (n = 110): | 21% |
| ²Local Lab IHC =3+ (n = 146): | 26% |

¹Intergroup Trial N9831; First-line Herceptin + other in node+/HER2+ pts J Clin Oncol 22:567, 2004 and 24:3032, 2006


*Consequence = Over Treatment*
HER2 Testing in NSABP-B31

“Inexperienced” vs. “Experienced” Labs

Central QC first 104 pts (JNCI 94:852, 2002)

<table>
<thead>
<tr>
<th>Outside Lab Results</th>
<th>Central Lab</th>
</tr>
</thead>
<tbody>
<tr>
<td>IHC=3+/Small labs (n=52):</td>
<td>IHC&lt;3+</td>
</tr>
<tr>
<td></td>
<td>19%</td>
</tr>
<tr>
<td>IHC=3+/Large labs (n=28):</td>
<td>FISH=Neg</td>
</tr>
<tr>
<td></td>
<td>-</td>
</tr>
<tr>
<td>FISH=Pos/All labs (n=27):</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>0%</td>
</tr>
</tbody>
</table>

19% false positives by IHC from “small” labs!
## HER2 Testing in NSABP-B31

### Results after QA Program

*Laboratory Approval Based on Large Volume and/or High Concordance IHC vs. FISH*

Central QC next 240 pts (BCRT 76: abst #9, 2002)

<table>
<thead>
<tr>
<th>Outside Lab Results</th>
<th>Central Lab</th>
</tr>
</thead>
<tbody>
<tr>
<td>IHC=3+ (n=104):</td>
<td>FISH=Neg</td>
</tr>
<tr>
<td></td>
<td>2%</td>
</tr>
</tbody>
</table>

False positives decreased from 19% to 2%.

**Experience/QA = Improved Performance!**
HER2 Testing by FISH

Orange = HER2
Green  = Chromosome
Ratio   = \#HER2 \div \#C17
Normal  = 1
Amp     > 1 (variable Hx)

**Problematic issues with FISH:**
Expensive  
**Difficult and time consuming**  
Diverse arbitrary cutoffs  
Errors (~10% overall; esp. low-level amplification)  
Other...
Studies Comparing IHC (HercepTest) vs. FISH (PathVysion)

<table>
<thead>
<tr>
<th>Study</th>
<th># Pts</th>
<th>0%</th>
<th>1+</th>
<th>2+</th>
<th>3+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Am J Clin Pathol 113:852, 2000</td>
<td>100</td>
<td>0%</td>
<td>0%</td>
<td>17%</td>
<td>89%</td>
</tr>
<tr>
<td>Proc ASCO 19:#294), 2000</td>
<td>142</td>
<td>0%</td>
<td>0%</td>
<td>31%</td>
<td>100%</td>
</tr>
<tr>
<td>Ann Oncol 12&quot;597, 2001</td>
<td>142</td>
<td>0%</td>
<td>0%</td>
<td>25%</td>
<td>100%</td>
</tr>
<tr>
<td>J Clin Oncol 19:354, 2001</td>
<td>79</td>
<td>-</td>
<td>-</td>
<td>25%</td>
<td>100%</td>
</tr>
<tr>
<td>J Clin Oncol 19:2714, 2001</td>
<td>145</td>
<td>-</td>
<td>-</td>
<td>13%</td>
<td>75%</td>
</tr>
<tr>
<td>JNCI 94:855, 2002</td>
<td>119</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>68%</td>
</tr>
<tr>
<td>J Pathol 199:418, 2003</td>
<td>426</td>
<td>0%</td>
<td>1%</td>
<td>48%</td>
<td>94%</td>
</tr>
</tbody>
</table>

~ 70% = IHC 0/1+ (0% FISH+)
~ 15% = IHC 2+ (30% FISH+)
~ 15% = IHC 3+ (95% FISH+)

~ 95% Overall Agreement
Non-Linear Correlation Between Gene Amplification and Protein Expression

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>#Genes/Ratio</th>
<th>#Receptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDA231</td>
<td>2.4/1.1</td>
<td>~20,000</td>
</tr>
<tr>
<td>MDA175</td>
<td>3.0/1.6</td>
<td>~90,000</td>
</tr>
<tr>
<td>MDA453</td>
<td>5.2/2.4</td>
<td>~500,000</td>
</tr>
<tr>
<td>SKBR3</td>
<td>15.3/4.5</td>
<td>~2,500,000</td>
</tr>
</tbody>
</table>

High Correlation Surprising!

6-Fold Increase
125-Fold Increase

Courtesy Dr. Ken Bloom
Clarient, Inc.
## Misconception: FISH > IHC

*Oncology Drug Advisory Committee (ODAC) data leading to FDA approval for Herceptin*

### IHC vs. FISH and Response to Herceptin

#### H0648 (Chemo ± Herceptin)

<table>
<thead>
<tr>
<th>Test Results</th>
<th>RR TTP</th>
<th>95% CI</th>
<th>#Pts</th>
</tr>
</thead>
<tbody>
<tr>
<td>3+</td>
<td>0.42</td>
<td>0.32-0.55</td>
<td>349</td>
</tr>
<tr>
<td>2+</td>
<td>0.82</td>
<td>0.54-1.24</td>
<td>120</td>
</tr>
<tr>
<td>FISH+</td>
<td>0.44</td>
<td>0.34-0.75</td>
<td>325</td>
</tr>
<tr>
<td>FISH-</td>
<td>0.66</td>
<td>0.45-0.99</td>
<td>126</td>
</tr>
</tbody>
</table>

#### H0649 (second/third line Rx)

<table>
<thead>
<tr>
<th>Test Results</th>
<th>RR TTP</th>
<th>95% CI</th>
<th>#Pts</th>
</tr>
</thead>
<tbody>
<tr>
<td>3+/FISH+</td>
<td>0.42</td>
<td>0.32-0.55</td>
<td>293</td>
</tr>
<tr>
<td>3+/FISH-</td>
<td>0.40</td>
<td>0.19-0.87</td>
<td>43</td>
</tr>
<tr>
<td>2+/FISH+</td>
<td>0.72</td>
<td>0.31-1.64</td>
<td>32</td>
</tr>
<tr>
<td>2+/FISH-</td>
<td>0.86</td>
<td>0.53-1.38</td>
<td>83</td>
</tr>
</tbody>
</table>

IHC≈FISH in H0649 (second/third line Rx)

IHC≈FISH in H0650 (first line Rx)
**Misconception: FISH > IHC**

*Oncology Drug Advisory Committee (ODAC) data leading to FDA approval for Herceptin*

IHC vs. FISH and Response to Herceptin

H0648 (Chemo ± Herceptin)

<table>
<thead>
<tr>
<th>Test Results</th>
<th>RR TTP</th>
<th>95% CI</th>
<th>#Pts</th>
</tr>
</thead>
<tbody>
<tr>
<td>3+/FISH+</td>
<td>0.42</td>
<td>0.32-0.55</td>
<td>293</td>
</tr>
<tr>
<td>3+/FISH-</td>
<td>0.82</td>
<td>0.54-1.24</td>
<td>120</td>
</tr>
<tr>
<td>2+/FISH+</td>
<td>0.44</td>
<td>0.34-0.75</td>
<td>325</td>
</tr>
<tr>
<td>2+/FISH-</td>
<td>0.66</td>
<td>0.45-0.99</td>
<td>126</td>
</tr>
</tbody>
</table>

```
Direct comparative statements of equivalence or superiority between IHC and FISH cannot be made.
```

Publications combined IHC 3+/2+ versus FISH+ ⇒ biased comparison and primary source of misconception that FISH better than IHC.

IHC≈FISH in H0650 (first line Rx)
ABSTRACT

Purpose: To develop a guideline to improve the accuracy of human epidermal growth factor receptor 2 (HER2) testing in invasive breast cancer and its utility as a predictive marker. Methods: The American Society of Clinical Oncology and the College of American Pathologists (CAP) convened an expert panel, which conducted a systematic review of the literature and developed recommendations for optimal HER2 testing performance. The guideline was reviewed by selected experts and approved by the board of directors for both organizations. Results: Approximately 20% of current HER2 testing may be inaccurate. When carefully validated testing is performed, available data do not clearly demonstrate the superiority of either immunohistochemistry (IHC) or in situ hybridization (ISH) as a predictor of benefit from anti-HER2 therapy. Recommendations: The panel recommends that HER2 status should be determined for all invasive breast cancer. A testing algorithm that relies on accurate, reproducible assay performance, including newly available types of brightfield ISH, is proposed. Elements to reliably reduce assay variation (for example, specimen handling, assay exclusion, and reporting criteria) are specified. An algorithm defining positive, equivocal, and negative values for both HER2 protein expression and gene amplification is recommended: a positive HER2 result is IHC staining of 3+ (uniform, intense membrane staining of >30% of invasive tumor cells), a fluorescent in situ hybridization (FISH) result of more than 6 HER2 gene copies per nucleus, or a FISH ratio (HER2 gene signals to chromosome 17 signals) of more than 2.2; a negative result is an IHC staining of 0 or 1+, a FISH result of less than 4.0 HER2 gene copies per nucleus, or a FISH ratio of less than 1.8. Equivocal results require additional action for final determination. It is recommended that to perform HER2 testing, laboratories show 95% concordance with another validated test for positive and negative assay values. The panel strongly recommends validation of laboratory assay or modifications, use of standardized operating procedures, and compliance with new testing criteria to be monitored with the use of stringent laboratory accreditation standards, proficiency testing, and competency assessment. The panel recommends that HER2 testing be done in a CAP-accredited laboratory or in a laboratory that meets the accreditation and proficiency testing requirements set out by this document.
2007 ASCO/CAP Guidelines for IHC Testing
J Clin Oncol 25:118 and Arch Pathol Lab Med 131:18

Assumes the use of accurate, standardized, and validated assays.

CLIA/CAP laboratory accreditation will require continuous comprehensive monitoring and validation.
Assumes the use of an accurate, standardized, and validated assays.

CLIA/CAP laboratory accreditation will require continuous comprehensive monitoring and validation.

**EXAs/LBAs...**

...taught us that ERα and PgR are weak prognostic factors in untreated patients but strong predictive factors for response to hormonal therapy.

**IHC replaced EXAs/LBAs on FFPE Samples.**

Is this good?


EXAs/LBAs...

>200 studies (technique, concordance, outcome)

Greene et al. Monoclonal antibodies to human estrogen receptor. PNAS 77:5115, 1980


Advantages of Assessing Hormone Receptors by IHC vs. LBA

**LBA**
- large specimens
- frozen specimens
- radioactivity
- difficult
- expensive
- signal any cells

**IHC**
- any size (small)
- frozen or fixed
- none (safe)
- easy
- cheap
- signal tumor cells

- many steps
  ⇒ not so easy
Evaluation of ER$\alpha$ and PgR is necessary in all primary breast cancers and both LBA and IHC are approved for routine clinical use (Category I).

Acknowledged persistent problems with IHC such as lack of standardization and validation...
Crisis in ER\(\alpha\) and PgR Testing in Breast Cancer by IHC

What is the Problem?

*Inaccurate Test Results*

Primarily False-Negatives

What are the Consequences?

*Improper Therapy*

Primarily Under-Treatment

Potentially Catastrophic

What is the Magnitude?

*Largely Unknown*

1000s of Tests each Day in 1000s of Laboratories Worldwide

Monitoring of Quality is Rare

Estimated at 10-20\% for ER\(\alpha\)...Higher for PgR?
**ERα and PgR Testing in Breast Cancer by IHC**

**Evidence and Magnitude of Problem**

*Personal Experiences of Health-Care Providers:*

*Largest and Most Convincing*

*Example: Re-testing (2\textsuperscript{nd} opinion) for ERα reveals \( \sim 30\% \) false-negative rate in consult practice of DC Allred.*

*News Reports and Scientific Publications of Uncovered Mistakes:*

*Rare but Compelling Examples:*

**BIG 1-98 Clinical Trail:** Predictive power of ERα and PgR by IHC (central vs. local labs) in receptor-positive, postmenopausal, early breast cancer randomized to adjuvant letrozole vs. tamoxifen.


69\% false-negatives for ERα by IHC (n\sim 100)!

44\% false-negative for PgR by IHC (n\sim 1,200)!

**Debacle in Newfoundland and Labrador, Canada**

40\% False-Negatives in \( \sim 2,000 \) originally ER-Negative Breast Cancers tested between 1997-2005!!!

Hede A. JNCI 100:836, 2008

**ERα and PgR Testing in Breast Cancer by IHC**

**Evidence and Magnitude of Problem**

**Published Results of Quality Assurance Programs:**

*Rare but Compelling Examples:*

**UK National External Quality Assessment Service (NEQAS).**

*Comprehensive proficiency assessment (accuracy/reproducibility) of testing for ERα by IHC in 150 laboratories in 26 countries worldwide.*

**ERα Test Cases**
- low pos (~25% cells)
- med pos (~50% cells)
- high pos (>75% cells)

**Each lab’s “in-house” assay**

**Reviewed by experts**

**Labs finding >10% pos cells**
- 40%
- 85%
- 99%

**~20% Overall Error**

**Antigen retrieval most problematic technical error.**

Example of Validating an IHC Assay for ERα

Harvey et al. J Clin Oncol 17:1474, 1999

- Fixed/archival samples
- Antibody 6F11 (ERα specific)
- Tris pH 9 HIER
- SAHRPx detection system
- H₂O₂/DAB/OsO₄ chromogen
- Methyl green counterstain

- Calibrated to clinical outcome
Choice of ERα Antibody
(pilot study 200 IBC by PS IHC)

% cases >0% pos cells: (sensitivity)

<table>
<thead>
<tr>
<th>Antibody</th>
<th>LBA</th>
<th>H222</th>
<th>1D5</th>
<th>LH2</th>
<th>6F11</th>
</tr>
</thead>
<tbody>
<tr>
<td>% cases &gt;0% pos cells</td>
<td>68</td>
<td>53</td>
<td>66</td>
<td>54</td>
<td>71</td>
</tr>
</tbody>
</table>

% agreement with LBA: (specificity)

- 70 79 68 87
Scoring Immunostained Slides

Proportion Score (PS) 0 → 1 → 1/100 → 2 → 1/10 → 3 → 1/3 → 4 → 2/3 → 5 → 1

Intensity Score (IS) 0 = negative 1 = weak 2 = intermed 3 = strong

Total Score (TS) = PS + IS (range 0-8)

Example of Scoring

Proportion Score = 4 (1/3rd to 2/3rds positive cells)
Intensity Score = 2 (average intensity “intermediate”)
Total Score = 6/8
Interpreting ER by IHC

Requires comprehensive studies calibrating content with response to hormonal therapy.

<table>
<thead>
<tr>
<th>IHC Score</th>
<th>LBA (fmol/mg)</th>
<th>#Patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>517 (26%)</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>67 (3%)</td>
</tr>
<tr>
<td>3</td>
<td>59</td>
<td>117 (6%)</td>
</tr>
<tr>
<td>4</td>
<td>67</td>
<td>190 (10%)</td>
</tr>
<tr>
<td>5</td>
<td>104</td>
<td>320 (16%)</td>
</tr>
<tr>
<td>6</td>
<td>141</td>
<td>370 (19%)</td>
</tr>
<tr>
<td>7</td>
<td>193</td>
<td>318 (16%)</td>
</tr>
<tr>
<td>8</td>
<td>282</td>
<td>83 (4%)</td>
</tr>
</tbody>
</table>

Near-Linear Correlation with LBA

Wide Distribution of Expression

Direct Correlation with Response to Tamoxifen

Positive >1% ER+ cells!
# ERα by IHC vs. LBA Predicting Clinical Outcome

<table>
<thead>
<tr>
<th></th>
<th>DFS</th>
<th></th>
<th>OS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>p-val</td>
<td>HR</td>
<td>p-val</td>
</tr>
<tr>
<td>No Rx (n=688)</td>
<td>IHC: 0.90</td>
<td>0.48</td>
<td>0.76</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>LBA: 0.74</td>
<td>0.62</td>
<td>0.79</td>
<td>0.0001</td>
</tr>
<tr>
<td>Chemo Only (n=404)</td>
<td>IHC: 1.01</td>
<td>0.96</td>
<td>0.97</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>LBA: 0.97</td>
<td>0.86</td>
<td>0.82</td>
<td>0.23</td>
</tr>
<tr>
<td>Chemo+Endo (Tam) (n=260)</td>
<td>IHC: 0.56</td>
<td>0.008</td>
<td>0.50</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>LBA: 0.51</td>
<td>0.01</td>
<td>0.58</td>
<td>0.05</td>
</tr>
<tr>
<td>Endo Only (Tam) (n = 517)</td>
<td>IHC: 0.47</td>
<td>0.0008</td>
<td>0.35</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>LBA: 0.71</td>
<td>0.32</td>
<td>0.38</td>
<td>0.0003</td>
</tr>
</tbody>
</table>
Additional Studies
Using Same/Similar IHC Assay for ERα

In Adjuvant Setting:

Love et al. J Clin Oncol 20:2559-2566, 2002 (OvEx ± Tam)
Horiguchi et al. Oncol Rep 14:1109, 2005 (Misc Endo Rx)

In Neoadjuvant Setting:

Ellis et al. J Clin Oncol 19:3808, 2001 (Tam vs. AI)
Ellis et al. JNCI 100:1380, 2008 (Tam vs. AI)

In Advanced Disease:

Elledge et al. Int J Cancer. 2000;89:111 (Tam)
Arpino et al. Clin Cancer Res 10:5670, 2004 (Tam)
Comprehensively Validated IHC Assays for Measuring Hormone Receptors in Breast Cancer

For ER\(\alpha\):

For PgR:
- Mohsin et al. Modern Pathol 17, 1545, 2004

Shortcoming of all is Absence of Untreated Cohorts to Separate Prognosis from Prediction...Relatively Minor Since Prognostic Power so Small?
Elements of Good Testing in Clinical Practice (Highly Inter-Related)
NOT Easy to do Right!

True Expertise (most important issue and largest problem)
Understand clinical utility (formal study; multidisciplinary conferences, etc.)
Understand technology (formal study; workshops; experience, etc.)
Don’t assume expertise just because you’re a “pathologist” (most with little training in IHC)

Technically Validated Assays
Follow a clinically validated assay
Standard operating procedures (don’t fiddle with the assay...)
Optimal sample preparation (e.g. 10% NBF 6-48 hours), identification, and tracking
Demonstrate and maintain evidence of adequate sensitivity, specificity, and reproducibility

Clinically Validated Assays
Technically validated assay calibrated to COMPREHENSIVE clinical outcome
Multiple randomized clinical trials (difficult)

Comprehensive Reports
Proper identification (patient, physician, times of collection-testing-reporting, etc.)
Quantitative results (proportion and intensity of positive cells)
Interpretation of results (positive vs. negative based on clinical benefit)
Critical ancillary information
Criteria for scoring and interpretation
Key references regarding clinical utility overall and of assay utilized
Key reagents (esp. primary antibody)
Assurance of ongoing quality assurance program

Quality Assurance Program
Uniform and stable scoring and interpretation (within and between pathologists)
Ongoing confirmation of competency
Uniform and stable quality of slides (especially appropriate positive and negative controls)
Accurate and stable distribution of results
Ongoing monitoring (reassurance that problems are identified and solved in time)
Internal and external confirmation of accuracy

Comprehensive Record Keeping