Clinical Practice Patterns and Cost Effectiveness of Human Epidermal Growth Receptor 2 Testing Strategies in Breast Cancer Patients

Kathryn A. Phillips, PhD; Deborah A. Marshall, PhD; Jennifer S. Haas, MD, MSPH; Elena B. Elkin, PhD; Su-Ying Liang, PhD; Michael J. Hassett, MD; Ilia Ferrusi, BSc; Jane E. Brock, MBBS, PhD; and Stephanie L. Van Bebber, MSc

BACKGROUND: Testing technologies are increasingly used to target cancer therapies. Human epidermal growth factor receptor 2 (HER2) testing to target trastuzumab for patients with breast cancer provides insights into the evidence needed for emerging testing technologies. METHODS: The authors reviewed literature on HER2 test utilization and cost effectiveness of HER2 testing for patients with breast cancer. They examined available evidence on: percentage of eligible patients tested for HER2; test methods used; concordance of test results between community and central/reference laboratories; use of trastuzumab by HER2 test result; and cost effectiveness of testing strategies. RESULTS: Little evidence was available to determine whether all eligible patients are tested, how many are retested to confirm results, and how many with negative HER2 test results still receive trastuzumab. Studies suggested that up to 66% of eligible patients had no documentation of testing in claims records, up to 20% of patients receiving trastuzumab were not tested or had no documentation of a positive test, and 20% of HER2 results may be incorrect. Few cost-effectiveness analyses of trastuzumab explicitly considered the economic implications of various testing strategies. CONCLUSIONS: There was little information about the actual use of HER2 testing in clinical practice, but evidence suggested important variations in testing practices and key gaps in knowledge.
Testing technologies have emerged as important components of healthcare and are accelerating rapidly for cancer therapies. There are already several well-known examples of the use of testing and molecular targeted therapies with more such therapeutics in the pipeline. These new testing technologies are essential ingredients in a shift towards personalized medicine (healthcare targeting medical interventions to patients based on their individual characteristics, particularly their genetics), and there is hope that they can improve health outcomes and reduce expenditures. A critical challenge to the implementation of targeted therapies in oncology is determining whether and how they will be provided to the individuals who will benefit most from them.

Human epidermal growth factor receptor 2 (HER2) testing to target trastuzumab treatment (Herceptin, Genentech, Inc., South San Francisco, Calif) for patients with breast cancer provides an instructive case study to inform discussion of the use of emerging testing technologies in clinical practice. HER2 testing was developed to determine which patients have breast cancers that overexpress the gene HER2, and for those 20%–30% of patients, treatment with trastuzumab, a humanized monoclonal antibody, proved to be highly effective. Without the ability to target the drug to this specific population, the drug would not have been approved. HER2 testing is a well-known example of the successful use of testing to target cancer treatment that has been used in clinical practice for older than 10 years. In 1998, the US Food and Drug Administration (FDA) approved trastuzumab and an accompanying test for use in patients with metastatic breast cancer. Their use was expanded to patients with early-stage breast cancer after 2005, and testing is now recommended for all patients with invasive breast cancer.

Our objective was to examine what is known about the utilization and cost effectiveness of HER2 testing in clinical practice in the United States. We restricted our examination to testing practices in the United States because practice patterns are likely influenced by healthcare coverage and reimbursement policies, which vary substantially across countries.

Despite the clinical success of trastuzumab, there is growing debate among patients, providers, and payers regarding the best methods for selecting patients for treatment based on HER2 test results. In particular, there is uncertainty regarding the most appropriate and efficient testing strategy and concerns about the reliability and interpretation of test results. This study provides an in-depth analysis and illustration of the broader issue of evidence gaps for new testing technologies and the issues emerging because of an increased emphasis on the need to translate basic science findings into clinical therapies and health benefits.

**Description of HER2 Testing**

The FDA has approved 3 types of tests to assess HER2 status in breast tissue: immunohistochemistry (IHC), fluorescence in situ hybridization (FISH), and, most recently, chromogen in situ hybridization (CISH). There are several commercially available test kits for both the IHC and the FISH tests with different performance characteristics, and there is no consensus about what the optimal methods are. Current guidelines from the American Society of Clinical Oncology (ASCO), the College of American Pathologists (CAP), and the National Comprehensive Cancer Network (NCCN) recommend using either IHC (with confirmation of indeterminate results by FISH) or FISH to determine HER2 status. Although FISH has been shown to be a better predictor of response to treatment, IHC is approximately one quarter to one third the cost of FISH and is more easily performed in community laboratories. Current guidelines state that a tumor with an IHC score of 3+, an average HER2 gene:chromosome 17 ratio of greater than 2.2 by FISH, or an average number of HER2 gene copies per cell of 6 or greater, is considered HER2 positive. A tumor with an IHC score of 2+ should be further tested using FISH, with HER2 status determined by the FISH result, although there is debate.
over the benefit of trastuzumab in patients with indeterminate FISH results. Testing decisions have recently become even more complex because of preliminary reports that women with test results that are negative based on recommended thresholds may actually benefit from trastuzumab. Thus, methods of evaluating eligibility for trastuzumab are still in flux.

Framework for Evaluating the Use and Cost Effectiveness of Testing

The effectiveness and efficiency of diagnostic testing in actual practice is based on a sequence of events: a test is ordered, test results are reviewed and interpreted to assess whether a patient is likely to benefit from a specific therapy, and treatment is offered and accepted or declined in accordance with test results (Fig. 1). Variation in these factors influences the actual effectiveness and cost effectiveness of testing and treatment.

Research Questions

1. Percentage of eligible patients tested for HER2 and methods used

Although there is clinical consensus that all patients with invasive breast cancer should be tested, some patients may not receive testing. In some cases patients may be tested, but their test results may not be properly documented.

We examined available evidence addressing the following questions:

- Do all patients with invasive breast cancer receive HER2 testing?
- Does test utilization vary according to personal characteristics, such as insurance status, socioeconomic status, and race/ethnicity, or by physician, health plan, or geographic characteristics?

We considered that patients may be tested by using IHC, FISH, or both methods but did not examine CISH because it has only recently been an FDA-approved option for clinical practice. Although there is no consensus about whether IHC or FISH should be used for initial testing, there is agreement that patients with indeterminate IHC results (approximately 15% of patients) should have a FISH test. However, some patients with indeterminate IHC results may not get confirmatory testing and there is continuing debate over the cutoff used. We examined available evidence addressing the following questions:

- What percentage of patients receives IHC, FISH, or both tests?
- Do patients with indeterminate IHC results get confirmatory FISH testing?
- Does the use of IHC or FISH vary by personal characteristics, such as insurance status, socioeconomic status, and race/ethnicity, or by physician, health plan, or geographic characteristics?

2. Percentage of tests with concordant results between community and central or reference laboratories

Even if positivity criteria can be determined so that only patients defined as having positive HER2 results are those who will benefit from trastuzumab, HER2 test performance in actual practice will be influenced by variability in laboratory procedures. Because the purpose of this study is to examine actual testing practices, our literature review focused on laboratory concordance rather than predetermined thresholds for determining positive results. We examined available evidence addressing the following question:

- What is the concordance between HER2 tests performed in community laboratories as compared with central or reference laboratories?

3. Percentage of patients who receive trastuzumab

Although clinical guidelines state that only patients testing HER2 positive according to established algorithms should receive trastuzumab, patients who either are not tested or have negative results may still receive trastuzumab because of doubts about test accuracy, clinical
factors, personal or healthcare system factors, or human error. Patients with positive HER2 test results may not receive trastuzumab because of the same factors. We examined available evidence addressing the following questions:

- What is the percentage of patients with positive, negative, or no test results who receive trastuzumab?
- Does the use of trastuzumab (conditional on test results) vary by personal characteristics such as insurance status, socioeconomic status, and race/ethnicity, or by physician, health plan, or geographic characteristics?

4. Cost effectiveness of HER2 testing in clinical practice

Studies that have examined the cost effectiveness of trastuzumab in a population of patients already identified as HER2 positive have found that adding it to chemotherapy is relatively cost-effective.10-12 These studies, however, do not address the cost effectiveness of different HER2 testing strategies in the adjuvant setting. For example, if patients who would benefit from treatment are not tested, if the most efficient testing algorithm is not used, if tests are inaccurate, or if patients receive treatment that is inappropriate based on their test results, then the cost effectiveness of testing and treatment in actual practice will be lower than the estimated cost effectiveness. We examined the available evidence on the following question:

- What is the cost effectiveness of HER2 testing strategies?

Literature Review

We conducted an extensive search to identify evidence on the HER2 testing practice patterns just described. To obtain data on actual use, we searched electronic databases, performed hand searches of relevant publications, did Internet searches, and reviewed key guidelines. Because much of this evidence is not in peer-reviewed publications due to its recency, we also included newspaper articles, abstracts, guidelines, and industry reports. To obtain data on cost effectiveness, we conducted a systematic search for cost effectiveness analyses of HER2 testing by using electronic databases and hand searching of relevant publications. For all searches, we included English-language publications through May 2008. Details of the search are provided in Table 1.

RESULTS

Evidence regarding testing practice patterns is summarized in Table 2 and described below.

Percentage of Eligible Patients Tested for HER2

Stark et al (2004), in a study of patients with metastatic breast cancer in the Henry Ford Health Care System (MI) shortly after the FDA approval of trastuzumab for patients with metastatic breast cancer (1999-2000), found that 52% were tested for HER2.13 Women with an absence of estrogen receptors, those with physicians in the surgery specialties, and those who had capitated insurance were more likely to be tested. A study of a 5% sample of Medicare enrollees in 2005 (N = 6588)—reported only in abstract form—showed that 32% of patients newly diagnosed with invasive breast cancer had documentation in claims data of having undergone a HER2 test.14 Of those with documentation of having received trastuzumab, 68% had documentation of having had a HER2 test. Among those receiving trastuzumab, older women and white women were more likely to have been tested.

Percentage of Patients Tested Using IHC and/or FISH Methods

The study noted above in the Medicare population also showed that 93% of the women tested received only IHC, 0.3% received only FISH, and 6% received both tests, although it is not known whether FISH was used for initial or confirmatory testing.14

Percentage of Tests With Concordant Results Between Community and Central or Reference Laboratories

We found no studies of laboratory concordance in routine practice, although several studies have examined concordance in conjunction with enrollment into clinical trials and can be presumed to reflect routine practice in general.15-18 The ASCO/CAP guidelines reviewed such studies and found that approximately 20% of IHC tests done
Table 1. Literature Review Methods*

<table>
<thead>
<tr>
<th>HER2 Testing Practice Patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The following search engines and electronic databases were queried:</td>
</tr>
<tr>
<td>- PubMed Database</td>
</tr>
<tr>
<td>- BIOSIS Previews Database</td>
</tr>
<tr>
<td>- International Pharmaceutical Abstracts Database</td>
</tr>
<tr>
<td>- National Cancer Institute Website</td>
</tr>
<tr>
<td>- Google Search Engine</td>
</tr>
<tr>
<td><strong>Search terms included:</strong></td>
</tr>
<tr>
<td>- Keywords: “HER2,” “erbB2,” “trastuzumab,” “Herceptin,” “use” or “utilization”</td>
</tr>
<tr>
<td>2. Popular medical and scientific peer-reviewed journals were manually searched: JAMA, New England Journal of Medicine, Health Services Research, Health Affairs, and Science</td>
</tr>
<tr>
<td>3. Abstracts presented at ASCO meetings were manually searched by authors</td>
</tr>
<tr>
<td>4. Guidelines released by ASCO/CAP and NCCN were manually searched by authors</td>
</tr>
</tbody>
</table>

Cost-Effectiveness Analyses of HER2 Testing or Treatment†

<table>
<thead>
<tr>
<th>Cost-Effectiveness Analyses of HER2 Testing or Treatment†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The following search engines and electronic databases were queried:</td>
</tr>
<tr>
<td>- PubMed</td>
</tr>
<tr>
<td>- BIOSIS</td>
</tr>
<tr>
<td>- Cochrane</td>
</tr>
<tr>
<td>- Centre for Reviews and Dissemination (CRD)</td>
</tr>
<tr>
<td>- EconLit</td>
</tr>
<tr>
<td>- EMBASE</td>
</tr>
<tr>
<td>- Health Economic Evaluations Database (HEED)</td>
</tr>
<tr>
<td><strong>MeSH search terms and other Keywords included:</strong></td>
</tr>
<tr>
<td>- economic (cost and cost-analysis, health economics, economic evaluation, pharmaconoeconomics, cost-effectiveness analysis, cost-benefit analysis, cost-utility analysis)</td>
</tr>
<tr>
<td>- breast cancer (breast tumor, breast carcinoma, breast neoplasm, mammary or breast and [tumor or carcinoma or neoplasm or cancer])</td>
</tr>
<tr>
<td>- trastuzumab or HER2 (ERB2 receptor, epidermal growth factor receptor, Herceptin, trastuzumab, HER2/neu)</td>
</tr>
<tr>
<td>2. Abstracts from the 2004-2006 San Antonio Breast Cancer Symposium were manually searched by authors</td>
</tr>
<tr>
<td>3. Technology appraisals from the National Institute for Health and Clinical Excellence (NICE) were manually searched by authors</td>
</tr>
</tbody>
</table>

* Abstracts and reports were excluded because these provided insufficient information for analysis.
† Criteria for inclusion: original research, cost-effectiveness analysis study, evaluation of HER2 testing strategies or trastuzumab treatment of HER2-positive breast cancer (early-stage or metastatic, and U.S. population.
by local, community-based laboratories are inaccurate when compared with central or reference laboratory results.7

**Percentage of Patients Testing Positive or Negative Who Received Trastuzumab**

One study of patients with early-stage or metastatic breast cancer in a privately insured population was conducted by using 2005 United HealthCare data (Lee Newcomer, oral and written communication, July 2008).19,20 They found that at least 12% and “probably” 20% of women receiving trastuzumab either did not have a HER2 test or had a test that showed no conclusive evidence of HER2 overexpression. More specifically, it was estimated that 8% were underexpressed, 4% had not been tested, and 8% were unknown because the physician did not provide the record. Details of this study have not been published in the peer-reviewed literature.

**Cost Effectiveness of HER2 Testing in Clinical Practice**

Of 621 studies screened, we found 4 cost-effectiveness analyses of HER2 testing or trastuzumab treatment in US populations.10-12,21 Two of these studies did not consider HER2 testing and examined only treatment strategies, assuming that the patients had already been tested.11,12 Another study included testing in their model but did not compare different testing strategies.10 One study examined HER2 testing strategies in a US population by analyzing 7 possible test-treat strategies for patients with metastatic breast cancer.21 We did not find any analyses that examined testing strategies in the adjuvant setting.

**DISCUSSION**

**Summary and Implications for HER2 Testing**

There is little evidence about the use of HER2 testing in routine clinical practice that can inform the current debate about selection of patients for treatment or the relative advantages and disadvantages of alternative testing strategies. The limited evidence available suggests that there are important variations in testing practices and key gaps in knowledge about those practices.

This study illustrates the gaps in what is known and the weaknesses in available evidence. Studies are generally of selected populations, use outdated data, and often are not published in peer-reviewed journals.

Understanding the use of HER2 testing in cancer care is important because of its critical role in targeting therapy. Withholding trastuzumab from a patient with HER2-positive breast cancer (underuse) or giving trastuzumab to a patient with HER2 negative breast cancer (overuse) will result in suboptimal outcomes. In the former situation, patients may not benefit from the substantial reduction in the risk of death from breast cancer conferred by trastuzumab.22 In the latter situation, patients who will not benefit from treatment are exposed to an unnecessary risk of heart failure and the healthcare system incurs unnecessary costs of about $100,000 annually for trastuzumab.7

---

**Table 2. HER2 Testing Practice Patterns**

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Summary of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of patients tested for HER2</td>
<td>- 52% of metastatic breast cancer patients at Henry Ford Health System (MI) in 1999/2000 (Stark 200413)</td>
</tr>
<tr>
<td>Percentage of patients tested by using IHC and/or FISH methods</td>
<td>- 32% of Medicare enrollees with newly diagnosed breast cancer in 2005 (Tong 200714)</td>
</tr>
<tr>
<td>Percentage of tests with concordant results between community and central or reference labs</td>
<td>- Among Medicare enrollees with metastatic breast cancer in 2005, 93% received only IHC, 0.3% received only FISH, 6% received both (Tong 200714)</td>
</tr>
<tr>
<td>Percentage of patients testing positive or negative who received trastuzumab</td>
<td>- Approximately 20% of HER2 tests performed by community laboratories are incorrect based on comparison to central or reference labs (Wolff 20077)</td>
</tr>
<tr>
<td></td>
<td>- 12%-20% of women who received trastuzumab in a large health plan had either not had the test or no conclusive evidence of a positive test (Renshaw 200619; Culliton 200820; Lee Newcomer, oral and written communication, July 18, 2008)</td>
</tr>
</tbody>
</table>

HER2 indicates human epidermal growth factor receptor 2; IHC, immunohistochemistry; FISH, fluorescence in situ hybridization.
Our results do not challenge the efficacy of HER2 testing and trastuzumab treatment; rather, they suggest a need to collect data on the utilization of tests and treatment and incorporate this information into rigorous analyses of best testing practices. It appears that some clinicians and payers assume that all eligible patients are being tested, tests are accurate, and only patients with positive test results are receiving trastuzumab. Our review suggests that gaps in the literature are substantial, and that these important assumptions cannot yet be verified. Although it is likely that testing has improved over time, this study demonstrates that even if one assumes that these problems are no longer issues, the paradigm outlined by HER2 and trastuzumab is likely to be repeated over and over as new, targeted therapies are developed. If we do not understand the mistakes that were made as this test-treatment combination was introduced, then we are likely to repeat them.

Testing is now recommended for all invasive breast cancers, but no studies examine how utilization has changed since 2005, when testing indications were expanded to include women with early stage breast cancers. It is particularly concerning that there are no studies of testing among typically underserved patients, including the uninsured, Medicaid insured, and minorities.

Little is known about how many patients receive 1 or both of the tests currently recommended and to what extent those patients are retested to confirm the results. As guidelines recommend using either IHC or FISH, it is important to know which method is used and how that usage varies by patient, provider, and healthcare system factors. In addition, considering the substantial percentage of patients who have indeterminate IHC results that require FISH testing for confirmation, it is important to know whether those patients actually get confirmatory testing. It is not known how the introduction of a third FDA-approved test (CISH) will improve or further widen the gap of evidence on tests used.

In contrast to the lack of evidence for our other research questions, there is more solid evidence to suggest that a substantial percentage of HER2 tests performed by community laboratories are inaccurate, based on comparisons with higher-volume central or reference laboratories. The available studies may not be representative of routine practice, however, because they were conducted as a part of clinical trials and may, therefore, underestimate the extent of problems with test accuracy. However, there may be times when community labs have more accurate results, eg, if the lesional tissue is small, then it may no longer be present on samples submitted to the reference lab or if the reference lab mistake DCIS for invasive carcinoma. However, we do not know of any studies that suggest the likelihood a patient will respond to the drug is more consistent with data from community versus central labs.

Reasons for discordance between laboratory results are complex and include differences in how laboratories perform and interpret tests as well as the lack of consensus about accepted procedures. Several initiatives are being pursued to resolve these issues and standardize laboratory performance, and there is increasing scrutiny of in house laboratory (“homebrew”) tests; for example, recently developed guidelines recommend that HER2 testing be done in a CAP-accredited laboratory. Considering the serious implications of inaccurate tests for patients’ lives and the impact on the healthcare system, it is essential to have more data on test quality and interpretation. As one industry observer noted, “We need more rigorous validation processes. It has been like the wild, wild West out there.”

It is also important to note that one study found that up to 20% of women receiving trastuzumab had no documentation of a positive HER2 test, and one study found that up to 66% of eligible patients had no documentation of testing in claims records. However, these findings have not been published in detail, are not based on recent data, and do not distinguish between women who were not tested and those who had been tested but were missing documentation of testing or results in claims records. We found similar results in a pilot study of one health plan, and we have a larger and more recent study underway (i3 Innovus, unpublished data, 2007). These findings argue for more standardized test reporting practices, with requirements such as the inclusion of both quantitative and qualitative assessments from pathology reports.

Actual practice patterns may have a substantial impact on the costs and effectiveness of testing and treatment in the adjuvant setting, yet such analyses have not been conducted. Analyses examining the costs and effectiveness of current testing guidelines and how they might vary in actual practice are sorely needed.

Implications

Our findings regarding HER2 testing illustrate both the challenges and the opportunities in building an evidence
base to support effective and efficient decision making about emerging testing technologies in cancer care. HER2 testing provides an example of a test that is clinically beneficial but that faces quality and implementation challenges, and such challenges will increasingly become relevant as more new testing technologies and targeted therapies emerge.24,25 The underlying science and associated guidelines for testing technologies are evolving quickly, and the different, new technologies make it harder to ensure that testing is done correctly and the results are interpreted appropriately. HER2 testing, as with many diagnostic tests, does not provide a black and white answer about treatment decisions.

With complex testing scenarios, there is ample room for mistakes and misinterpretation along the entire testing sequence.23 Thus, one solution is greater standardization of test procedures and processes. Communication between the laboratory and the clinician may be of particular importance; for example, in the case of HER2 testing, the speculation is that patients with indeterminate scores are being treated as positive results.23 Decision analytic tools for clinicians could facilitate improved interpretations and communications.

Another solution is increasing the amount of data available on testing technologies. Historically, there have been less data available on diagnostics than on other interventions, such as pharmaceuticals or surgical procedures, a gap that has taken on greater significance as the use of diagnostics to target therapies has accelerated. One major gap is the lack of administrative databases linking testing, test results, treatment, and outcomes. In the case of HER2 testing, claims databases do not typically include information on test results. Moreover, it is often impossible to identify the use of testing in administrative databases because of a lack of billing codes specific to each test or type of test received by a patient. In health insurance claims, the use of IHC and FISH tests for HER2 detection cannot be reliably distinguished from the same types of assays performed for other indications. Although Current Procedural Terminology code modifiers have been developed by the American Medical Association to differentiate specific tests, these modifiers are not commonly used in clinical practice. In addition, test codes may be bundled into a common pathology code that does not permit the identification of individual tests.

Conclusions

The trend toward greater use of testing to target healthcare is inevitable and has the potential to improve the quality and efficiency of healthcare. Our findings should not be construed as reasons for attempting to slow the diffusion of new testing technologies into clinical practice and cancer care or as a critique on oncologists’ treatment decisions. Rather, this case study highlights how and what evidence might be improved to help guide decisions regarding emerging tests and associated therapies in cancer care. Given the rapid growth in this area—for example, there are more than 6000 articles on gene-disease associations each year and more than 1300 genetic tests making their way to market—, evidence-based information will become a necessity if these new technologies are to be used wisely.26

It is crucial to build an evidence base that can support effective and efficient decision making in regard to emerging testing technologies in clinical practice, considering their impact on clinical care and the healthcare system.27 A comprehensive agenda for translational research is needed to move new discoveries, such as those in human genomics, into health practice in a manner that maximizes health benefits and minimizes harm to individual people and populations.6,26 This agenda will require increased attention to the translation of not only basic research findings into new therapeutic options but also research into actual practice.5 By examining translational issues, we can improve the effectiveness and efficiency of care, reduce disparities, and ensure that the patients who will most benefit from treatment will get it.

Conflict of Interest Disclosures

This study was funded by 2 grants to Dr. Phillips from the National Cancer Institute (R01CA101849 and P01CA130818) and a grant from the Blue Shield Foundation of California (unrestricted). Funding agencies had no review of any portion of this work.

Drs. Phillips and Haas have an unrestricted grant from the Aetna Foundation to examine the utilization of HER2 testing and gene expression profiling for breast cancer. Dr. Haas has a research grant from Pfizer for unrelated work (on adverse drug event reporting).

References


