Centre universitaire de santé McGill
McGill University Health Centre

FINAL

Report No 23

TESTING FOR HER2 POSITIVE BREAST CANCER: A COST- EFFECTIVENESS ANALYSIS

Report available at

www.mcgill.ca/tau/

May 15th 2006

This report was prepared for the Technology Assessment Unit (TAU)

of the McGill University Health Centre (MUHC)

by

Nandini Dendukuri, PhD , Research Scientist, Technology Assessment Unit
James Brophy, Divisions of Cardiology and Clinical Epidemiology, Department of Medicine, MUHC

and approved and adopted by the committee of the TAU:
Juliana Arnoldo, Jeffrey Barkun, Andre Bonnici, James Brophy, Pierre Ernst, John Johnston, Marilyn Kaplow, Maurice McGregor, Gary Pekeles, Judith Ritchie, Gary Stoopler, Donatella Tampieri.

Acknowledgments

The expert assistance of the following individuals is gratefully acknowledged:

Consultants:

Karim Khetani
(Department of Pathology)

Invitation.

This document was developed to assist decision-making in the McGill University Health Centre. All are welcome to make use of it. However, to help us estimate its impact, it would be deeply appreciated if potential users could inform us whether it has influenced policy decisions in any way.

E-mail address:

maurice.mcgregor@mcgill.ca ***james.brophy@mcgill.ca***

Table of Contents

EXECUTIVE SUMMARY

Background The need for accurate detection of HER2 positive status among women with breast cancer has come into focus due to the availability of a targeted therapy (trastuzumab or Herceptin) for this condition. The two most commonly used diagnostic tests for detection of HER2 positive status are: 1) immunohistochemistry (IHC) that detects over-expression of the HER2 protein, and 2) fluorescence in-situ hybridization (FISH) that detects amplification of the HER2 gene. Though IHC is significantly less expensive and easier to perform than FISH, there are concerns that it is less accurate and less reliable than FISH. Further, it has been found that FISH positive women respond to trastuzumab treatment while FISH negative women do not. The current practice at the McGill University Health Centre (MUHC) and elsewhere is to consider an IHC score of 3+ as HER2 positive, IHC scores of 0 or 1+ as HER2 negative and an IHC score of 2+ as ambiguous. The IHC 2+ category is re-tested by FISH.

Objective: 1) To systematically review the literature on the validity and reliability of IHC and FISH. 2) To carry out a cost-effectiveness analysis to compare various testing strategies based on IHC and FISH.

Methods: We carried out a systematic literature search using the PUBMED online database. Articles were selected if they satisfied the following inclusion criteria: 1) published in English in a peer-reviewed journal, 2) included human subjects, 3) reported the validity of IHC and FISH with respect to a superior standard, or reported the concordance between FISH and IHC, or the reliability of either test using standard cut-offs, 4) used assays licensed by Health Canada for identifying patients eligible for trastuzumab treatment and followed manufacturers instructions, 5) had a sample size of at least 100.

FISH was considered the gold-standard test. The validity of IHC was defined as the probability of a FISH positive result in each IHC score category. The reliability of IHC or FISH was defined as the percentage agreement between repeated tests carried out either by the same/different observers/laboratories. A meta-analysis was carried out to estimate the distribution of IHC scores and to estimate the probability of positive FISH results conditional on IHC scores, i.e. to determine IHC validity. Only studies that did not have a selection bias were used in the estimation of the distribution of IHC scores.

A Bayesian cost-effectiveness analysis was carried out to compare different testing strategies to the current strategy at the MUHC in terms of the incremental cost per accurate diagnosis. The current cost of IHC is $108 per case compared to $467 per case for FISH. The cost of 1 year of treatment with trastuzumab is estimated at $50,000 per year. The annual cost of new equipment needed to perform FISH at the MUHC is estimated at $5,000. The number of new breast cancer patients at the MUHC each year is estimated to be 320.

Results: A total of 18 studies satisfied the inclusion criteria. Most studies evaluated the PathVysion™ FISH assay and HercepTest™ IHC assay. The percentage of patients in each IHC score category (median) was estimated as - IHC 0: 36.1, IHC 1+: 35.5, IHC 2+: 12.2, IHC 3+: 16.2. The percentage of positive FISH results (median) among the 4 IHC categories was estimated as - IHC 0: 1.7, IHC 1+: 3.4, IHC 2+: 29.9, 3+: 91.9. In general, studies of reliability between/within observers/laboratories, concluded that the reliability of FISH was better than for IHC. However, there was evidence that FISH neither had 100% sensitivity and specificity nor 100% reliability. The currently pursued strategy of confirmatory testing for 2+ cases alone was expected to correctly diagnose 310 of 320 (96.9%) women. Among strategies that result in an increase in the number of correctly diagnosed cases, the strategy of confirmatory testing for 2+ and 3+ cases is associated with the lowest incremental cost-effectiveness ratio (ICER) of $5,784/accurately diagnosed case compared to current strategy. The strategy of performing FISH on all cases is associated with the highest ICER of $9,445/accurately diagnosed case. Our sensitivity analysis revealed that if the cost of FISH decreases to below $200 per test a strategy of testing all women with FISH may become the most cost-effective. Also if new equipment needs to be purchased for performing FISH at the MUHC, the strategy of testing all women with FISH becomes more cost-effective relative to a strategy of confirmatory testing for patients with 1+, 2+ and 3+ scores.

Limitations: This study has not analyzed the cost-effectiveness of trastuzumab treatment. This evaluation of HER2 testing is predicated on a decision that trastuzumab treatment is accepted.

Conclusions: With the current arrangement for FISH testing external to the MUHC, the most cost-effective strategy is to screen all patients with IHC, followed by confirmatory

testing with FISH of those patients with IHC scores of 2+ or 3+. Purchase of new equipment may be justified if it is decided to test all patients with FISH.

Recommendation: If it is assumed that trastuzumab therapy will be offered, it is recommended that all breast cancer cases be screened with IHC and those who have scores of 2+ or 3+ be tested by FISH to confirm their HER2 positive status. In the unusual event that a breast cancer patient approaches the hospital with a positive FISH test carried out outside the MUHC, an IHC test needs to be carried out to confirm the course of action.

GLOSSARY

CHF: Congestive Heart Failure
CI: Confidence Interval obtained using Frequentist statistical inference
CrI: Credible Interval obtained using Bayesian statistical inference
FISH: Fluorescence In Situ Hybridization
HER2: Human Epidermal growth factor Receptor 2
HercepTest™: An immunohistochemistry assay licensed by Health Canada for identifying candidates for trastuzumab treatment.
ICER: Incremental Cost-Effectiveness Ratio
IHC: ImmunoHistoChemistry assay
LVEF: Left Ventricular Ejection Fraction
MUHC: McGill University Health Centre
PathVysion™: A fluorescence in situ hybridization assay licensed by Health Canada for identifying candidates for trastuzumab treatment.
Pathway™ An immunohistochemistry assay licensed by Health Canada for identifying candidates for trastuzumab treatment.
PharmDx™: A fluorescence in situ hybridization assay licensed by Health Canada for identifying candidates for trastuzumab treatment.
Polysomy: Having one or a few chromosomes present in a greater number than is characteristic of the rest of the chromosome complement
RAMQ: Régie de l'assurance Maladie du Québec
QALY : Quality-adjusted Life Year

INTRODUCTION

HER2 positive breast cancer is characterized by amplification of the HER2/neu (human epidermal receptor-2) or c-erbB-2 oncogene and/or overexpression of its protein[1]. In a normal cell producing HER2, there are two copies of the gene and about 50,000 copies of the protein at the cell surface. In comparison, in HER2 producing cancers there are more than 2 copies of the gene and more than 1,000,000 copies of the protein at the cell surface. HER-2 positivity occurs in 20-30% of breast cancers. It has been associated with more rapid tumor growth, increased risk of recurrence following surgery, shortened survival and poor response to conventional chemotherapy[2]. Thus detection of HER2 status has become an integral part of the clinical-pathological work-up of breast cancer[3].

The availability of a promising targeted monoclonal antibody therapy for HER2 positive breast cancer, trastuzumab (Herceptin), has focused further interest on the accurate detection of HER2 status [4-8]. Despite the reported efficacy of trastuzumab, it is associated with deleterious side-effects and a high cost of roughly $50,000 per year[9-11]. HER2 status has also been shown to predict response to adjuvant doxorubicin chemotherapy[12]. Thus it is important to minimize the number of both false-negative as well as false-positive results. Trastuzumab was initially approved for women with metastatic cancer but recently it has also been approved for use in localized breast cancer[10]. This increases the importance of accurately testing for HER2 as it implies that all breast cancer cases may need to be tested at diagnosis.

Currently the two most widely used methods for detection of HER2 positive status are immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH). Both tests can be carried out in routinely collected formalin-fixed paraffin-embedded clinical samples[1]. While IHC is directed at detecting protein overexpression at the cell membrane, FISH is directed towards detection of gene-amplification. As of October 2005, Health Canada had licensed two FISH assays (PathVysion™ manufactured by Abbott Laboratories, pharmDx™ manufactured by DAKO Cytomation) and two immunohistochemistry assays (Hercep Test™ manufactured by DAKO Cytomation, Pathway™ manufactured by Ventana Medical Systems) for identifying women eligible for trastuzumab treatment. Though other FISH and immunohistochemistry assays (e.g. the Ventana INFORM™ FISH assay and the Ventana CONFIRM™

immunohistochemistry assay) have been licensed to determine HER2 status for prognostic purposes, they have not been licensed for the purpose of identifying patients eligible for trastuzumab treatment.

The absence of a gold-standard test for HER2, together with reported differences in reliability and cost of IHC and FISH, has fueled an intense debate on which is the better test to use[13]. Although FISH is generally believed to be the better test, the high concordance between IHC and FISH has lent support to a testing strategy where patients are screened with IHC and only ambiguous results are re-tested with FISH[13;14]. However, studies suggesting that trastuzumab is beneficial only in patients who are FISH positive have raised the issue of whether this testing strategy is cost-effective[15-17].

In this report we summarize the literature on licensed IHC and FISH assays. Assuming FISH is the gold-standard, we estimate the number of false-positive and false-negative diagnoses under various testing strategies. We estimate the cost-effectiveness (cost per accurate diagnosis) of each testing strategy, taking into account the cost of subsequent trastuzumab treatment. This analysis should not be confused with a complete cost-effectiveness analysis of trastuzumab at different stages of breast cancer that takes into account its clinical efficacy and cost, as well as the cost-effectiveness of diagnostic testing.

Diagnostic testing for HER2 status

Immunohistochemistry (IHC):

Method:

Immunohistochemistry assays involve detection of antigens in tissue using polyclonal or monoclonal antibodies. Commercially available IHC assays are based on several different antibodies. The HercepTest™ is based on the A0485 polyclonal antibody, while the Pathway™ is based on the CB11 monoclonal antibody. A specially developed clinical trial assay (CTA) for selection of women into clinical trials of trastuzumab was based on the 4D5 and CB11 monoclonal antibodies.

Scoring:

The IHC test involves a semi-quantitative scoring method. Tumor samples are scored on the extent of circumferential membrane staining on a scale from 0-3[1]. When using the HercepTest™, those with no staining (scored 0) or partial staining in more than 10% of cells (scored 1+) are considered to be normal. Those with intense circumferential thick membrane staining (scored 3+) are considered HER2 positive and most likely to benefit from trastuzumab treatment. Tumor samples that demonstrate thin circumferential staining in more than 10% of cases (scored 2+) are considered ambiguous. The interpretation of the categories is slightly different when using the Pathway™ assay: 0 (No staining), 1+ (Faint, partial staining of the membrane), 2+ (Weak complete staining of the membrane in >10% of cancer cells), 3+ (Intense complete staining of the membrane in >10% of cancer cells).

Re-testing only the ambiguous results on IHC (i.e. the 2+ scores) by FISH is the currently recommended standard of practice[10;14].

Advantages:

The main advantage of IHC is that it is widely available in most surgical pathology laboratories. It is also much less expensive than FISH (estimated cost of $108 per test at the MUHC).

Disadvantages:

Pre-analytical variables such as tissue-handling and fixation can affect immunoreactivity. Further, the subjective nature of the scoring method leads to greater inter-observer variability. An automated image analysis system may address the latter problem. The availability of commercially manufactured assays is believed to have addressed the problem of standardization to some extent.

Fluorescence in situ hybridization (FISH):

Method:

In FISH testing a fluorescent-labeled probe is used to enumerate the HER2 gene copy number[1]. When the probe is added to the tissue section it recognizes and hybridizes with the target gene. Probe signals are required to be present in approximately 75% or more of cancer cell nuclei. The occurrence of polysomy of chromosome 17 can also be a cause of

gene amplification, though it occurs in a very small percentage of cases (<1%)[1]. Both assays licensed by Health Canada correct the HER2 copy number for the chromosome 17 copy number.

Scoring:

Scoring is in terms of the ratio of the number of HER2 signals to the number of chromosome 17 centromere signals (HER2/CEN17 ratio) per nucleus. A ratio greater than 2 is taken to be indicative of HER2 gene amplification. It is recommended to re-test borderline results of 1.8-2.2[1].

Advantages:

DNA is a more stable target than the HER2 protein and is less susceptible to problems associated with tissue handling or fixation.

Disadvantages:

It is expensive (estimated cost of $467 per test), time-consuming and requires special training. It has a higher failure rate than the IHC. For example, a recent study reported a failure rate of 5% for the PathVysion™ compared to 0.08% for the HercepTest™ [3]. Currently, the cost of repeating a FISH test is not borne by the MUHC.

***Prognostic significance of HER2: IHC vs FISH*:**

In a study aimed at comparing the association between overall survival predicted by IHC (based on the R60 antibody) and FISH, Pauletti et al[18] found significantly worse survival among women with a score of IHC 3+ compared to those with IHC scores of 0, 1+ or 2+. There was no significant difference in survival between the IHC 0-1+ and 2+ categories. When women with an IHC score of 0-1+ were further classified as positive (N=76) or negative (N=633) on FISH, there was no difference in the overall survival between the two sub-groups (p-value of log-rank test 0.2794). This would suggest that women with an IHC score of 0-1+ had a similar prognosis irrespective of their FISH result. However, among women with IHC 2+ and 3+ scores the overall survival was worse in the FISH positive group (N=102) than in the FISH negative group (N=45) (p-value of log-rank test 0.0044).

Trastuzumab

Trastuzumab is a directed therapy for women with HER2 positive breast cancer[1]. It is a monoclonal antibody that seeks out and binds to the specific HER2 or HER2/neu receptors on the surface of HER2 over-expressing breast cancer cells and directly inhibits tumor cell growth. Trastuzumab was first approved in 1998 for use in women with metastatic breast cancer[7].

Recently published results from three Phase III randomized controlled trials have demonstrated a beneficial effect of trastuzumab in women with localized invasive breast cancer[7;8]. In the study by Romond et al, data from two different clinical trials with similar designs were jointly analyzed[7]. Women were selected for this study if they had a score of 3+ on an IHC assay designed specifically for the trial, or if they had a FISH+ result. Following concerns with reliability of diagnostic tests carried out at outside facilities all cases were tested by the IHC assay at a central facility. It was found that the percentage of patients alive and disease-free at the end of 3 years was 75.4% in the control group and 87.1% in the trastuzumab group (Absolute difference: 11.7% (95% confidence interval: 8.1%, 15.4%))[7]. In the study by Piccart-Gebhart et al, patients were entered into the study if they had an IHC score of 3+ on HercepTest+™ or FISH+. The percentage of patients alive and disease-free at the end of 2 years was 77.4% in the control group and 85.8% in the treatment group (Absolute difference: 8.4% (95% confidence interval: 2.1%, 14.8%))[8].These results have lead to the approval in Quebec of trastuzumab for cases with node positive breast cancer or node negative breast cancer with tumor size exceeding 1cm, i.e. women with Stage II cancer and some with high risk Stage I cancer[9;10]. These women make up about 95% of breast cancer cases[19].

The enthusiasm surrounding trastuzumab has been balanced by a number of concerns. It has been associated with several undesirable side-effects, particularly high cardiac toxicity[7-9]. In both recently published clinical trials a large percentage of patients discontinued treatment – 31.4%[7] and 8.5%[8] - primarily due to adverse cardiac effects. In the study by Piccart-Gebhart et al, in the trastuzumab group 9/1677 (0.54%) experienced severe congestive heart failure (CHF), and 113/1677 (7.08%) experienced a drop in left ventricular ejection fraction (LVEF) of greater than 10%. In comparison in the control group 0/1710 (0%) had CHF and 34/1710 (2.21%) of patients experienced a drop in

LVEF[8]. In the study by Romond et al, in the trastuzumab group the cumulative incidence of cardiac events (CHF or death) over a 3-year period was 4.0% compared to 0.6% in the control group[7]. Current guidelines recommend treatment until disease recurrence for patients with metastatic disease and treatment for 1-year in patients with Stage I or II breast cancer[10]. Treatment costs are substantial at an estimated $50,000 per patient-year.

Efficacy of trastuzumab in HER2 positive women – IHC vs FISH

Using data from the three pivotal clinical trials of Herceptin in metastatic cancer[4-6], Mass et al conducted a retrospective analysis on the relation between FISH results and the efficacy of trastuzumab[15]. Women were selected for these trials if they had a score of 2+ or 3+ on a specially designed IHC assay called the Clinical Trials Assay (CTA). All women were later tested using the PathVysion™ FISH test. It was found that 169/765 (22%) of the selected patients for the trial were FISH negative. Of those who were FISH negative 122/169 (72%) had an IHC score of 2+. The authors compared groups that received chemotherapy+trastuzumab to chemotherapy alone within strata defined by FISH. The outcome variable was the objective response rate, defined as a 50% or greater reduction in the dimensions of all measurable lesions. Using data from one of the trials (H0648g) they found that there was no effect of the intervention on the objective response rate among the FISH negative strata (19/50 (38%) in the intervention group vs. 22/57 (39%) in the control group). However, there was a significant benefit of the intervention in the FISH positive strata (95/176 (54%) in the intervention group vs. 51/168 (30%) in the control group). In the remaining two trials also a better response rate to trastuzumab was observed in FISH positive than in FISH negative patients (H0649g: 33/173 (19%) vs 0/36 (0%), H0650g: 28/82 (34%) vs 2/29 (7%)). There was also a greater beneficial effect of trastuzumab among FISH positive patients for other outcomes such as time to disease progression and overall survival. No articles have been published yet on the relation between the results on the commercially available IHC assays and the efficacy of trastuzumab.

Cost effectiveness of trastuzumab

The high cost of trastuzumab treatment has emphasized the need for cost-effectiveness analyses. Several studies have estimated cost-effectiveness of trastuzumab treatment and are reviewed in this section. All studies were carried out assuming patients had metastatic cancer. All cost figures in this section are in Canadian dollars for ease of comparison. The actual figures reported in each study are in the footnote.

Only one study modeled different testing strategies using IHC and FISH[17]. The mean costs of IHC testing, FISH testing and 3-week cycle trastuzumab treatment were assumed to be $99, $445 and $2,687[i], respectively. Taking into account the quality-adjusted life-years (QALYs) following various treatment regimens they concluded it was more cost-effective to receive trastuzumab than to not receive it, if a patient was HER2 positive. Further, they concluded that the most cost-effective strategy was to identify eligible patients by first screening with IHC and then re-testing IHC 2+ and 3+ cases with FISH (Incremental Cost-Effectiveness Ratio (ICER) compared to a base case of no testing for HER2 and no treatment with trastuzumab: $146,226/QALY[ii]). Compared to a strategy of confirmatory testing for only IHC 2+ cases, their recommended strategy had a lower cost and no reduction in QALYs.

Other studies evaluating cost-effectiveness of trastuzumab have typically ignored the possibility of different testing strategies. Norum et al assumed that IHC 3+ would serve as the criterion for detecting HER2 positive patients[20], while Neyt et al assumed FISH testing would be required for all patients as per guidelines in Belgium[21]. Norum et al concluded that trastuzumab treatment was not cost-effective (ICER ranging from $86,814-223,321/life-year saved[iii])[20]. Neyt et al cautioned that the cost-effectiveness ratio was very high and that cost-effectiveness studies need to be carried out before introducing trastuzumab in different settings (ICER: $5,474 per life-month gained[iv])[21].

A report by the National Institutes for Clinical Excellence (NICE) in the United Kingdom reported cost-effectiveness of using trastuzumab together with paclitaxel based on data obtained from the manufacturer of trastuzumab (Roche)[22]. After including cost

[i] US$85, US$381 and US$2,301
[ii] US$125,000/QALY
[iii] €63,137-162,417/life-year saved
[iv] €3,981.44/life-month gained

of one-time IHC testing and cardiac testing four times, the ICER was $72,271/QALY[v][23]. The NICE report recommended that trastuzumab be used in combination with paclitaxel for women with an IHC score of 3+ who had not had chemotherapy and for whom anthracycline treatment was not appropriate. Trastuzumab monotherapy was recommended in women with an IHC score of 3+ who had at least two previous chemotherapy treatments

A Canadian study by Dranitsaris et al evaluated the timing of testing for HER2. They concluded that it was more cost-effective to test women at the time of diagnosis than at the time when the disease enters the metastatic phase[19].

The situation at the MUHC

Currently, all breast cancer cases are tested for HER2 status using the Pathway™ IHC assay. The estimated cost of each test is $108 (Test cost: $73 (MSSS[vi] code 60570) + Pathologist charge: $35 (RAMQ[vii] billing code: 10111)). Patients who receive scores of 3+ are considered eligible for treatment with trastuzumab. Tumor samples with scores of 2+ are sent to the Jewish General Hospital for FISH testing with the PathVysion™ assay. The estimated cost of each FISH test is $467 (Test cost: $422 (MSSS code 50719) + Pathologist charge: $45 (RAMQ billing code: 10101)). The current turnaround time is 2-5 days for IHC and 1-5 weeks for FISH. The MUHC is currently evaluating the exact cost of performing FISH in-house. It is anticipated that the cost could drop to below $400/test while the turn-around-time for FISH would drop to 1-2 weeks. However, if the volume of FISH testing increases, for example because a more stringent standard for confirmatory testing of IHC is employed, new equipment may need to be purchased.

From July 2004 to June 2005 a total of 304 breast biopsies from new or recurrent cases of breast cancer were tested for HER2. The IHC results for these patients are summarized in Table 1.

[v] £37,500/QALY
[vi] MSSS: Ministére de la Santé et des Services Sociaux
[vii] RAMQ: Régie de l'assurance Maladie du Québec

Table1: Distribution of IHC results at the MUHC

IHC Score	0	1+	2+	3+
Number (%)	159 (52.3%)	70 (23.0%)	41 (13.4%)	34 (11.1%)

In addition to this there are about 10-15 cases of suspected metastatic breast cancer for which HER2 testing was performed on biopsies from other anatomical sites.

METHODS

Search strategy for review:

We searched the PUBMED database using the following keywords: (sensitivity OR specificity OR reliability OR reproducibility OR validity OR interobserver OR intraobserver OR accuracy OR "predictive value") AND IHC AND FISH AND HER2 AND breast. Articles were selected if they satisfied the following inclusion criteria: 1) published in English or French in a peer-reviewed journal, 2) included human subjects, 3) reported either the concordance between FISH and IHC or the reliability of either test using standard cut-offs, 4) used assays licensed by Health Canada for identifying patients eligible for trastuzumab treatment and followed manufacturers instructions, 5) a sample size of at least 100. It has been reported that the IHC test has poor reliability, particularly in small laboratories[24]. Hence we limited our review to larger studies. We used an arbitrarily selected minimum sample size of 100. We also searched the bibliographies of articles identified by the PUBMED search.

Diagnostic tests are typically evaluated in terms of validity (by comparison to a reference standard) and reliability (comparison of results of multiple repetitions of a test). There is no perfect gold-standard test for HER2. However, since FISH is widely considered the better test we evaluated the validity of IHC by comparison with FISH. From articles on concordance between IHC and FISH, for each IHC score category we calculated:

FISH positivity rate in IHC score category x =

$$\frac{\text{Number of FISH positives in IHC score category x}}{\text{Number of patients in IHC score category x}},$$

where IHC score category x refers to the score categories 0, 1+, 2+ and 3+. Where possible we reported results separately for 0 and 1+ categories. We would expect the

lowest FISH positivity rate in the IHC 0 category, increasing to the highest FISH positivity rate in the IHC 3+ category.

Reliability may be measured between multiple test readings by the same or different observers/laboratories. From articles on reliability of either IHC or FISH we extracted information on concordance (or percentage agreement) between repeated measurements on the same test defined as:

$$\frac{\text{Number of patients who received the same result from both observers (positive or negative)}}{\text{Number of patients evaluated by both observers}} \times 100$$

The concordance ranges from 0% (no agreement) to 100% (perfect agreement). If the concordance was not available we extracted any other measure of reliability, such as the kappa coefficient. The kappa coefficient is a measure of the agreement between raters after having adjusted for agreement by chance[25]. It ranges from 0 (agreement no better than chance) to 1 (perfect agreement).

Meta-analysis:

A Bayesian meta-analysis was carried out to estimate the distribution of IHC scores and to estimate the probability of a positive FISH result in each IHC score category. The goal of this analysis was to estimate the proportion of accurately classified individuals based on IHC scores. The analysis was carried out in two parts. Studies that were known or suspected to have a selection bias were not included in the estimation of the distribution of IHC scores, but only included in the estimation of the FISH positivity rate. We required the patient sample to have a representative distribution of test results to limit verification bias. For example, having an over-representation of 2+ and 3+ individuals in the sample would artificially raise the sensitivity of the IHC and decrease its specificity[26]. In some cases it was apparent from the study methods that there was an over-representation of IHC 2+ or 3+ cases. When the percentage of IHC 2+ and 3+ cases exceeded 40% we treated the study as having selection bias even if this was not evident from the methods. We included in the analysis both studies that reported IHC 0 and IHC 1+ cases jointly or separately. Non-informative prior distributions were used for all parameters.

FISH was assumed to be a gold-standard. The analysis was implemented using the WinBUGS software program. After verifying convergence of the Gibbs Sampler, we used a burn-in of 1000 iterations and reported results using the next 5000 iterations. We repeated the analysis in the sub-set of studies where both the IHC test and FISH test were carried out at the same site. This was to eliminate studies where there may have been a greater variability in IHC performance due to the test being performed at multiple labs[27;28]. Also, in these studies we could only ascertain that the majority of IHC tests (but not all tests) had been carried out using a licensed assay.

Cost-effectiveness analysis:

We carried out a Bayesian cost-effectiveness analysis that was based on the assumption that trastuzumab is efficacious only among HER2 positive women. Thus, our interest was in comparing the cost-effectiveness of different strategies for identifying the group of women who would benefit from trastuzumab treatment.

The seven strategies that were compared are summarized in Table 2.

Table 2: Candidate diagnostic testing strategies for HER2 positive breast cancer.

Strategy	HER2 positivity criterion
1. IHC for all patients -> FISH for IHC 2+ (Base case)	IHC 3+ or FISH+
2. IHC for all patients	IHC 2+ or IHC 3+
3. IHC for all patients	IHC 3+
4. IHC for all patients -> FISH for IHC 1+, 2+	IHC 3+ or FISH+
5. IHC for all patients -> FISH for IHC 2+, 3+	FISH+
6. IHC for all patients -> FISH for IHC 1+, 2+, 3+	FISH+
7. FISH for all patients	FISH+

Costs of testing and treatment were obtained in consultation with experts at the MUHC. The current cost of IHC testing is $108. The current cost of FISH testing is $467. The probability of an accurate diagnosis based on each strategy was estimated using the posterior distributions obtained from the meta-analysis. The number of breast cancer

cases tested annually at the MUHC was assumed to be 320. We also estimated the cost-effectiveness of strategies 4-7 assuming purchase of new equipment by the MUHC for FISH testing. The analysis was carried out using the WinBUGS software package. The cost of the equipment was assumed to be $50,000 amortized over 10 years.

The current practice at the MUHC was termed the base case to which all remaining strategies were compared. Among strategies other than the base case, any strategy whose median cost was higher and whose median accuracy was lower than another strategy was considered to be dominated, and was eliminated from the analysis. Any strategy that had a lower median number of accurate diagnoses compared to the base case was also eliminated given that it is unlikely to be acceptable to lower the number of women whose HER2 positive status is accurately diagnosed. The remaining strategies were then compared to the base case using an incremental cost-effectiveness ratio.

Sensitivity analysis:

We assessed the robustness of our analysis by varying the input variables over the ranges summarized in Table 3. We carried out both a univariate sensitivity analysis (examining the change in ICER for competing strategies when changing one variable at a time), and a multivariate sensitivity analysis (examining the change in ICER for competing strategies when allowing all variables to change simultaneously). The purpose of the univariate analysis was to describe the effect of each variable on the ICERs of competing strategies. The purpose of the multivariate sensitivity analysis was to estimate the confidence interval around the ICERs when allowing for uncertainty in all the variables in our model.

The distributions of the IHC score and the FISH positivity rate was determined from the meta-analysis. For the cost variables we used a uniform distribution over the range of values of interest. For the multivariate analysis this range was defined as +/- 20% around the median value. Results were summarized as scatter plots of the ICERs vs each variable using the R software package.

Table 3: Range of variables used in sensitivity analyses

Variable	Univariate analysis *Median (Range)*	Multivariate analysis *Median (Range)*
IHC distribution (%)		
0	43.1 (4.7, 63..8)	43.1 (4.7, 63..8)
1+	35.5 (6.9, 66.7)	35.5 (6.9, 66.7)
2+	12.2 (3.4, 21.2)	12.2 (3.4, 21.2)
3+	16.2 (10.7, 22.9)	16.2 (10.7, 22.9)
FISH positivity rate (%)		
Among IHC 0	1.7 (0.9, 2.8)	1.7 (0.9, 2.8)
Among IHC 1+	3.4 (2.3, 5.1)	3.4 (2.3, 5.1)
Among IHC 2+	29.9 (11.2, 59.3)	29.9 (11.2, 59.3)
Among IHC 3+	91.9 (85.9, 95.9)	91.9 (85.9, 95.9)
Cost of IHC testing (Ca$)	108 (50, 150)	108 (86, 130)
Cost of FISH testing (Ca$)	467 (100, 600)	467 (374, 560)
Annual cost of new equipment for FISH testing (Ca$/year)	5,000 (1000, 10,000)	5,000 (4000, 6,000)

RESULTS

Articles satisfying inclusion criteria

The flow chart in Figure 1 summarizes the study selection process. We identified 18 studies that satisfied our inclusion criteria. Of the selected studies 5 reported both concordance between IHC and FISH and reliability of one or both tests[27-31], 11 reported concordance alone[3;32-41] and 2 reported reliability alone[24;42]. Only 2 of these studies compared IHC and FISH with an external standard. Most selected studies used PathVysion™ and HercepTest™. The Pathway™ test was evaluated by one study[38]. We did not find results on the pharmDx™ test in a peer-reviewed journal.

Sensitivity and specificity of IHC and FISH

Comparison with external standard:

The two studies which compared IHC and FISH with a third reference standard, that is considered more accurate, but not possible to use in the clinical setting, are summarized below. Both studies show that FISH is the more sensitive test (Tables 4 and 5). However, FISH is neither 100% sensitive nor specific.

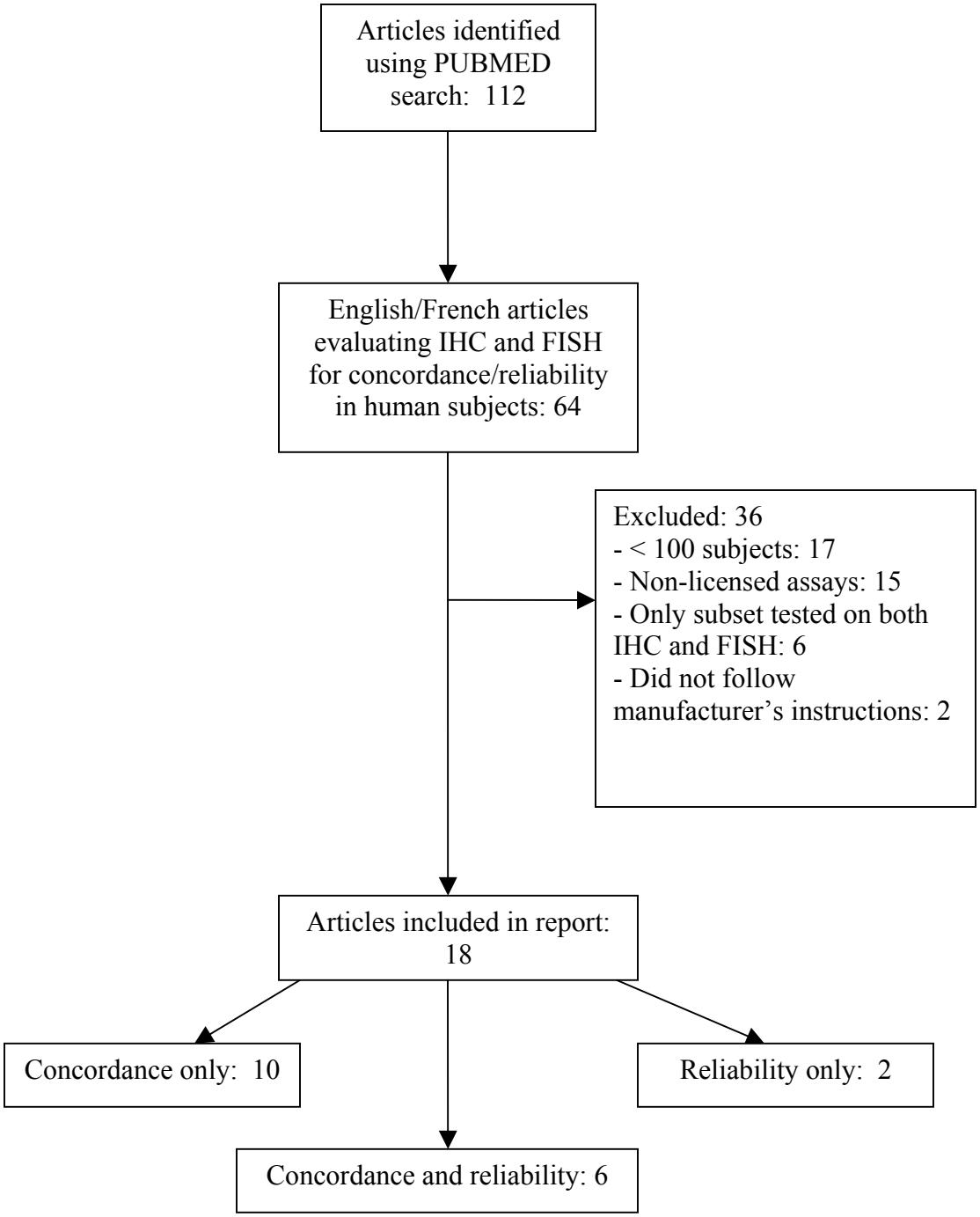
Bartlett et al, 2001[29]: Quantitative radio-immunohistochemistry (Q-IHC) was defined as the reference standard. A cut-off of 10 times or more overexpression than normal was considered indicative of positive HER2 status. The sample consisted of 191 frozen and formalin-fixed breast carcinoma sections. IHC was carried out using HercepTest™ and FISH was carried out using PathVysion™. Samples were independently scored by two blinded pathologists. The results are summarized in Table 4.

Table 4: Results on IHC and FISH sensitivity and specificity from Bartlett et al, 2001

	Sensitivity (N)	Specificity (N)
FISH (positive if HER:CEP17>=2)	84.1% (44)	94.6% (146)
IHC (positive if >= 2)	61.9% (42)	98.6% (143)

They concluded that while HercepTest™ was a reliable guide to therapy, its low sensitivity means it would miss a significant number of HER2 positive cases.

Figure 1: Flowchart of search strategy.



Press et al, 2002[32]: Solid matrix blotting was used as the reference standard. The sample consisted of 117 archival, paraffin-embedded specimens. Specimens were required to show agreement on both amplification and overexpression analyses by the reference standard. Results comparing HercepTest™ and PathVysion™ to the reference standard are summarized in Table 5.

Table 5: Results on IHC and FISH sensitivity and specificity from Press et al, 2002

	Sensitivity (N)	Specificity (N)
FISH (positive if HER:CEP17>=2)	95.4% (43)	98.6% (74)
IHC (positive if >= 2)	69.8% (43)	100% (74)
IHC (positive if >= 3)	39.5% (43)	100% (74)

The authors also found that three other IHC assays using different antibodies (R60 Polyclonal antibody, 10H8 Monoclonal Antibody, CB11 Monoclonal Antibody) had better overall accuracy than HercepTest™. They reported that the sensitivity of all tests was associated with the level of gene amplification. HercepTest™ correctly identified 9 out of 10 samples with more than 20-fold amplification, however it correctly identified only 7 out of 16 samples in which the amplification was between 2-5 fold. They concluded that FISH has a higher sensitivity than IHC.

Comparison between IHC and FISH

Summary of observed data:

Table 6 summarizes the results from studies that have evaluated the concordance between IHC and FISH. More details on the individual studies (sample selection process, tests used and conclusions) are given in the Appendix. Among studies that did not have selection bias (studies 1 – 7, 10 - 12), the percentage of IHC 0-1+ results ranged from 56.4% to 85.2%. From the subset of studies reporting IHC 1+ results separately (studies 10-12), the percentage of patients in this category varied from 8.4% to 47.3%. The percentage of IHC 2+ scores ranged from 2.0% to 15.5% and the percentage of IHC 3+ results ranged from 4.8% to 28.2%. Among studies that were likely to have selection bias (studies 8-9, 13-16) the percentage of at least one of the IHC 2+ or 3+ categories was

much higher. The FISH positivity rate within IHC categories ranged from 0% to 14.9% among IHC 0,1+, from 0% to 100% among IHC 2+, and from 66.4% to 100% among IHC 3+, across all studies. Among the subset of studies with information on IHC 1+, the FISH positivity rate ranged from 0% to 8.5%. Thus the greatest heterogeneity in FISH positivity was in the IHC 2+ category.

Results of Bayesian meta-analysis:

The results of the Bayesian meta-analysis for estimating the joint distribution of IHC and FISH are summarized in Table 7. Similar results were obtained in the subset of 14 studies where the IHC and FISH tests were performed at the same site. Therefore, we do not present these results separately. Based on our meta-analysis, the expected number (and 95% Credible Interval (CrI)) of FISH positive patients at the MUHC is 66 (49, 85). This means there is a 95% probability that the annual number of FISH positive patients at the MUHC is between 49 and 85, assuming the number of women tested is 320.

Table 6: Studies used in meta-analysis

Studies with results reported jointly for IHC 0 and 1+ categories							
Author (Year)	**N**	**Distribution of IHC results (%)**			**FISH positivity rate (%) by IHC score**		
		0,1+	**2+**	**3+**	**0,1+**	**2+**	**3+**
1. Lottner (2005)	215	78.1	11.6	10.2	2.4	72.0	100.0
2. Loring (2005)	110	56.4	15.5	28.2	0.0	0.0	87.1
3. Dowsett (2003)	426	63.4	12.7	23.9	0.7	48.1	94.1
4. Press (2002)	117	74.4	11.1	14.5	14.9	100.0	100.0
5. Bartlett (2001)	210	85.2	10.0	4.8	6.7	90.5	90.0
6. Hoang (2000)	100	74.0	2.0	24.0	0.0	0.0	70.8
7. Kakar (2000)	112	70.5	15.2	14.3	1.3	3.5	87.5
8. Mrozkowiak (2004) †	360	2.8	87.5	9.7	0	20.3	91.4
9. Yaziji (2004) †	2913	49.0	39.5	11.5	2.8	17.0	91.6
Studies with results reported separately for IHC 0 and 1+ categories								
Author (Year)	**N**	**Distribution of IHC results (%)**				**FISH positivity rate (%) by IHC scores**			
		0	**1+**	**2+**	**3+**	**0**	**1+**	**2+**	**3+**
10. Lal (2004)	2279	44.6	31.4	13.7	10.3	1.1	3.1	26.5	89.7
11. Ogura (2003)	110	24.6	47.3	9.1	18.2	3.7	3.8	10.0	100.0
12. Tsuda (2001)	101	41.6	34.7	5.9	17.8	2.4	2.9	0.0	83.3
13. McCormick (2002)†	198	26.8	29.8	22.7	20.7	3.8	8.5	3.5	87.5
14. Roche (2002) **, †	119	7.6	8.4	10.1	73.9	0	0	0	89.8
15. Press (2005) **,†	842	36.5	17.8	36.5	9.3	3.6	5.3	26.9	66.4
16. Dolan (2005) †	129	1.6	16.3	72.1	10.1	0	0	8.1	62.5

** HercepTest™ conducted at various laboratories compared with FISH at a central facility. † Distribution of sample test scores not representative of population.

Table 7: Results of meta-analysis.

		Results of Meta-Analysis (%)		Expected Numbers at the MUHC+**	
	IHC Result	Median	95% CrI	Median	95% CrI
Distribution of IHC results	0	36.1	4.7, 63.8	116	15, 204
	1+	35.5	6.9, 66.7	114	22, 213
	2+	12.2	3.4, 21.2	39	11, 68
	3+	16.2	10.7, 22.9	52	34, 73
FISH positive results	0	1.7	0.9, 2.8	2	0, 4
	1+	3.4	2.3, 5.1	4	1, 8
	2+	29.9	11.2, 59.3	12	2, 28
	3+	91.9	85.9, 95.9	48	31, 67

** Rounded to the nearest integer. Assuming 320 breast cancer patients/year at the MUHC.

Studies on reliability of IHC and FISH

A summary of studies on reliability of IHC and FISH is given in Table 8. Concordance on HercepTest^{TM} (IHC) varied from 75% to 91.8% across studies using various designs. While reliability of PathVysion^{TM} (FISH) was better in general, varying from 92.4% to 98.7%, it was not perfect. More details on the sample selection process and conclusions of individual studies are given in the Appendix.

Table 8: Summary of studies on reliability of IHC and FISH

First Author (Year)	Test		Reliability
	Type	Commercial Name	
1. Bartlett (2001)	IHC FISH	HercepTest™ PathVysion™	kappa = 0.67 kappa = 0.97
2. Paik (2001)	IHC	Various IHC assays (77% HercepTest™)	Concordance with central HercepTest™ = 82/104 (72.2%)
3. Roche (2002)	IHC	HercepTest™	44/59 (75%)
4. Dowsett (2003)	IHC	HercepTest™	150/180 (83.3%)
5. Dybdal (2005)	IHC	HercepTest™	448/488 (91.8%)
6. Press (2005)	FISH	FISH* (outside vs central) PathVysion™ (intra-laboratory) PathVysion™ (inter-laboratory)	121/131 (92.4%) 57/59 (96.6%) 75/76 (98.7%)
7. Hoang (2000)	IHC	HercepTest™	85/100 (85%)
8. Tsuda (2001)	FISH	PathVysion™	213/216 (99%)

* Not known which FISH assays were used outside. PathVysion™ used at central laboratory.

Cost-effectiveness analysis

The results of the cost-effectiveness analysis are presented in Table 9. All cost figures were rounded to the nearest dollar for simplicity. The estimated numbers of correctly and incorrectly classified patients following each strategy (rounded to the nearest integer) are presented in Table 10.

Under the base case strategy (Strategy 1) the cost of testing was estimated to be $52,840 and the number of accurately identified cases was estimated to be 310/320 (96.9%) (Table 9). Of the remaining strategies, the three that were not dominated or eliminated because of lower effectiveness were: Strategy 5) confirmatory testing for IHC 2+ and 3+ cases, Strategy 6) confirmatory testing for IHC 1+, 2+ and 3+, and Strategy 7) testing all patients with FISH.

Median ICERs for these strategies ranged from $5,784 (Strategy 5) to $9,445 (Strategy 7) (Table 9 and Figure 2). Under Strategy 5, the additional 4 correctly classified cases would all have been falsely classified as positive under the base case strategy. Thus this strategy would result in an estimated saving of about $200,000 in the cost of Herceptin treatment (Table 9). Strategies 6 and 7 are significantly more expensive as they involve testing the larger group of IHC 1+ patients who have a low FISH positivity rate. Testing all patients with FISH (Strategy 7) will result in an estimated annual diagnostic cost of $149,440, but will result in detection of approximately 6 additional cases that would be missed by Strategy 5 (Table 9).

If new equipment were to be purchased at an annual cost of $5,000, Strategy 5 would continue to be the most cost-effective strategy (Table 9 and Figure 2), however Strategy 7, would become more cost-effective than Strategy 6.

Sensitivity analysis

The results of the univariate sensitivity analysis are presented in plots in Figures 3a-3e. It appears that the cost of FISH is the only variable that could have an important effect on the ordering of the ICERS. As the cost of FISH drops to below about $200, Strategy 7 becomes the most cost effective.

In Figure 3a we see that Strategy 5 has a lower ICER than all strategies even if the expected number of individuals in the IHC 1+ category increased or decreased. However, as the percentage of individuals in the IHC 1+ category increase beyond 0.3 Strategy 7 becomes more cost-effective than Strategy 6.

In Figure 3b, we see that as the FISH positivity rate in the IHC 1+ category increases, Strategies 6 and 7 have decreasing ICERs. However, Strategy 5 remains the most cost-effective strategy across the range of possible values of this variable.

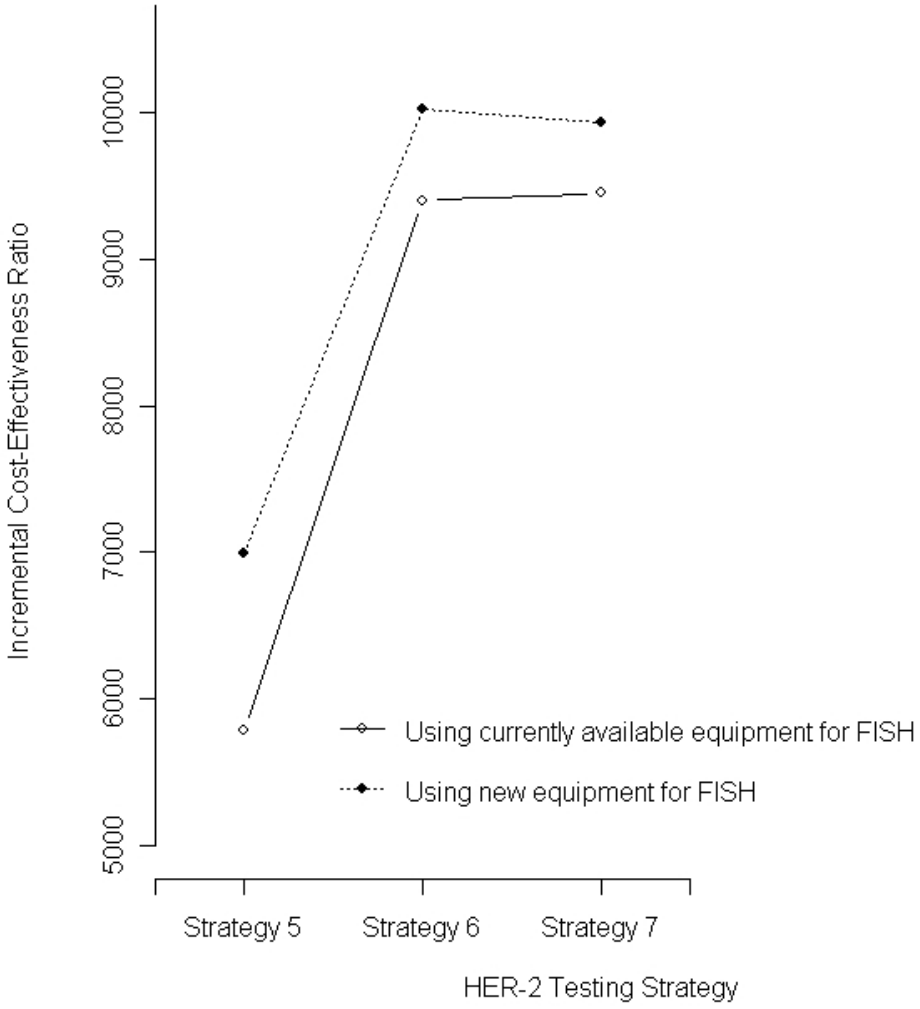
In Figure 3c we see that as the cost of FISH drops to below about $200 Strategy 7 becomes the most cost-effective.

In Figure 3d we see that as the cost of IHC increases beyond about $110 the cost-effectiveness of Strategy 7 decreases. However, Strategy 5 continues to have the best cost-effectiveness.

In Figure 3e, we see that as the annual cost of new equipment increases the ICERs of all strategies increase. Strategies 6 and 7 have similar ICERs, though as the cost of the equipment increases, Strategy 7 is more cost-effective.

From the multivariate sensitivity analysis we determined that the 95% confidence intervals for the ICERs ($/accurate diagnosis) of the three competing strategies using existing equipment for the FISH test were (3154; 11,440) for Strategy 5, (5495; 14,710) and (5,737; 15,110). If new equipment for FISH is to be purchased, the 95% confidence intervals for the ICERs ($/accurate diagnosis) were: (3,832; 13,730) for Strategy 5, (6,118; 15,530) for Strategy 6 and (6,104; 15,770) for Strategy 7.

Figure 2: Results of cost-effectiveness analysis.



Strategy 5. IHC for all patients -> FISH for IHC 2+, 3+ (Positivity criterion: FISH+)
Strategy 6. IHC for all patients -> FISH for IHC 1+, 2+, 3+ ((Positivity criterion: FISH+)
Strategy 7. FISH for all patients (Positivity criterion: FISH+)

Note: The lines on this plot are meant as a visual guide and do not imply a continuum of strategies between those listed on the x-axis.

Table 9: Results of cost effectiveness analysis (median (95% Credible intervals))

Strategy	HER2 positivity criterion	Number of FISH tests	Annual Cost ($)	Number accurately diagnosed (%)	ICER ($/accurate diagnosis)	Cost of Herceptin (in million $)
Using existing system for FISH testing						
1. IHC for all patients -> FISH for IHC 2+ (Base case)	***IHC 3+ or FISH+***	39 (10, 68)	52,840 (39,440; 66,460)	310 (305, 313)	-	3.2 (2.3, 4.2)
2. IHC for all patients	IHC 2+ or IHC 3+	0 (-)	34,560	284 (260, 302)	Dominated	2,6 (1.7, 3.7)
3. IHC for all patients	IHC 3+	0 (-)	34,560	299 (282, 307)	Eliminated	3.2 (2.3, 4.2)
4. IHC for all patients -> FISH for IHC 1+, 2+	IHC 3+ or FISH+	268 (246, 289)	106,000 (63, 460; 152, 300)	313 (310, 317)	Dominated	3.4 (2.5, 4.4)
5. IHC for all patients -> FISH for IHC 2+, 3+ (selected strategy)	***FISH+***	91 (70, 116)	77,090 (67,120; 88,540)	314 (311, 316)	5, 784 (3,303; 11,430)	3.0 (2.2, 4.0)

Strategy	HER2 positive criterion	Number of FISH tests	Annual cost ($)	Number accurately diagnosed (%)	ICER ($/accurate diagnosis)	Cost of Herceptin (in million $)
6. IHC for all patients -> FISH for IHC 1+, 2+, 3+	FISH+	205 (115, 306)	130,400 (88,310; 177,300)	318 (316, 320)	9,395 (5,838; 14,180)	3.2 (2.4, 4.2)
7. FISH for all patients	FISH+	320 (-)	149,440 (-)	320 (-)	9,445 (6,629; 13,890)	3.3 (2.5, 4.3)
Using new equipment for FISH testing ($5,000/year for 5 years)						
1.IHC for all patients -> FISH for IHC 2+ (base case)	***IHC 3+ or FISH+***	39 (10, 68)	57,840 (44,440; 71,460)	310 (305, 313)	-	3.2 (2.3, 4.2)
4. IHC for all patients -> FISH for IHC 0-1+, 2+	IHC 3+ or FISH+	268 (246, 289)	111,000 (68,460; 157,300)	313 (310, 317)	Dominated	3.4 (2.5, 4.4)
5. IHC for all patients -> FISH for IHC 2+, 3+ (selected strategy)	FISH+	91 (70, 116)	82,090 (72,120; 93,540)	314 (311, 316)	6,998 (4,002; 13,880)	3.0 (2.2, 4.0)
6. IHC for all patients -> FISH for IHC 1+, 2+, 3+	FISH+	205 (115, 306)	130,400 (93,310; 182,300)	318 (316, 320)	10,020 (6,381; 15,060)	3.2 (2.4, 4.2)
7. FISH for all patients	FISH+	320 (-)	154,440 (-)	320 (-)	9,931 (6,973; 14,610)	3.3 (2.5, 4.3)

Table 10: Distribution of true positive and false-positive results following each strategy

Strategy	HER2 positivity criterion	Classification				Number who receive Herceptin
		TP	FN	TN	FP	
1. IHC for all patients -> FISH for IHC 2+ (Base case)	***IHC 3+ or FISH+***	60 (43, 79)	6 (4, 9)	250 (230, 268)	4 (2, 8)	64 (47, 84)
2. IHC for all patients	IHC 2+ or IHC 3+	60 (43, 79)	6 (4, 9)	223 (199, 244)	30 (11, 54)	91 (69, 116)
3. IHC for all patients	IHC 3+	47 (31, 68)	17 (8, 34)	250 (230, 268)	4 (2, 8)	52 (34, 73)
4. IHC for all patients -> FISH for IHC 1+, 2+	IHC 3+ or FISH+	64 (47, 83)	2 (0, 4)	250 (230, 268)	4 (2, 8)	68 (50, 88)
5. IHC for all patients -> FISH for IHC 2+, 3+ (selected strategy)	***FISH+***	60 (43, 79)	6 (4, 9)	254 (235, 271)	0	60 (43, 79)
6. IHC for all patients -> FISH for IHC 1+, 2+, 3+	FISH+	64 (47, 83)	2 (0, 4)	254 (235, 271)	0	64 (47, 83)
7. FISH for all patients	FISH+	66 (49, 85)	0	254 (235, 271)	0	66 (49, 85)

Figure 3a: Relation between ICERs and probability of IHC 1+

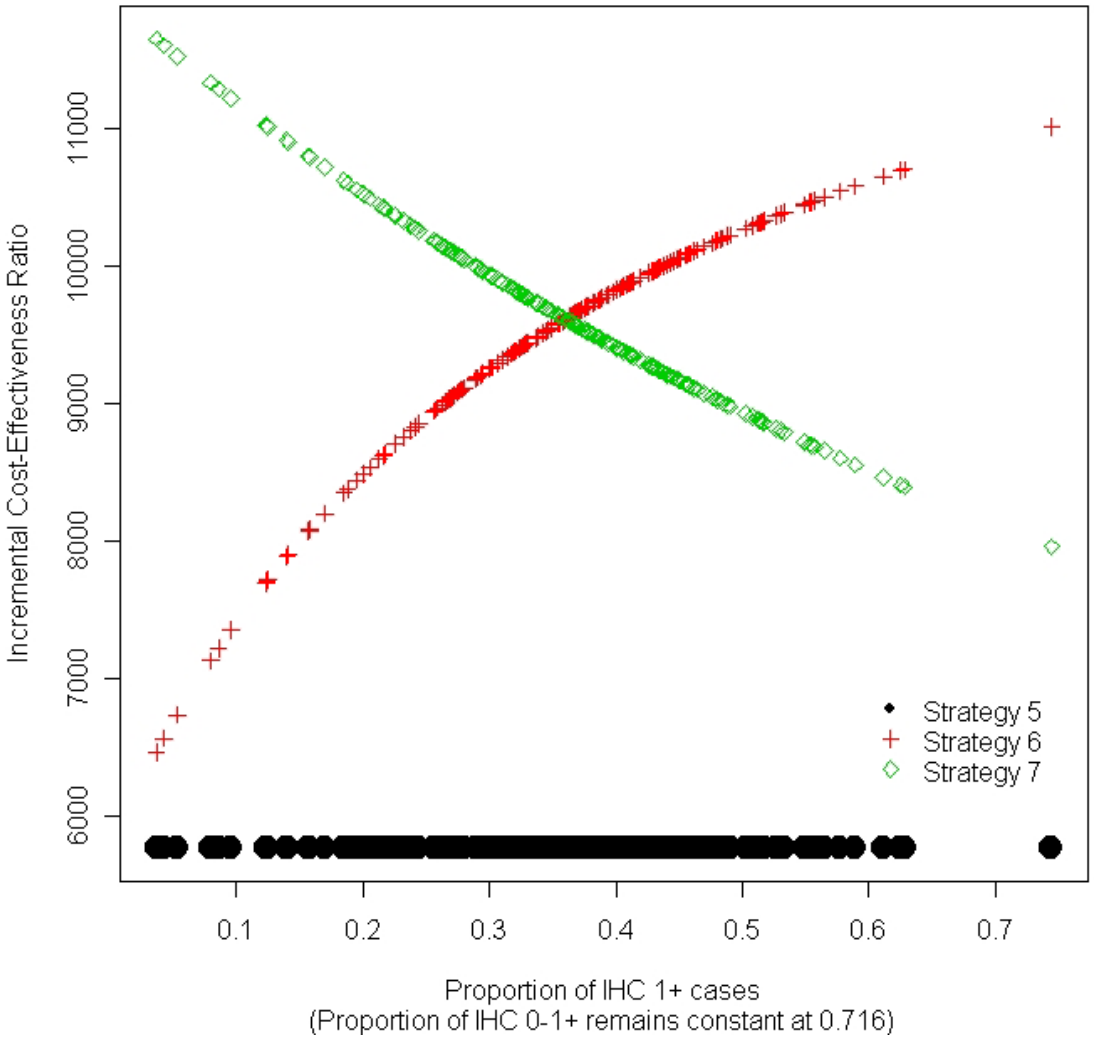


Figure 3b: Relation between ICERs and FISH positivity rate in IHC 1+

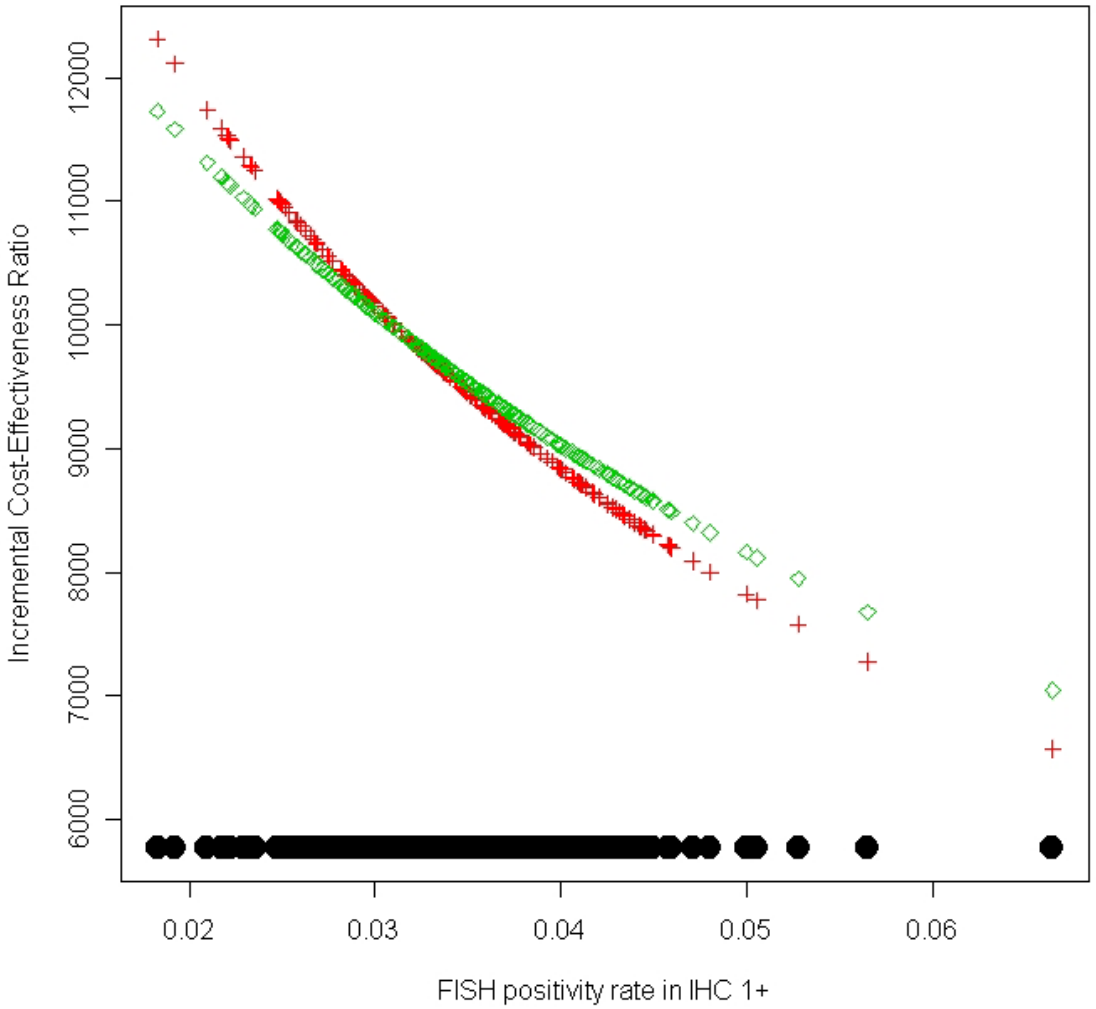


Figure 3c: Relation between ICER and cost of IHC*

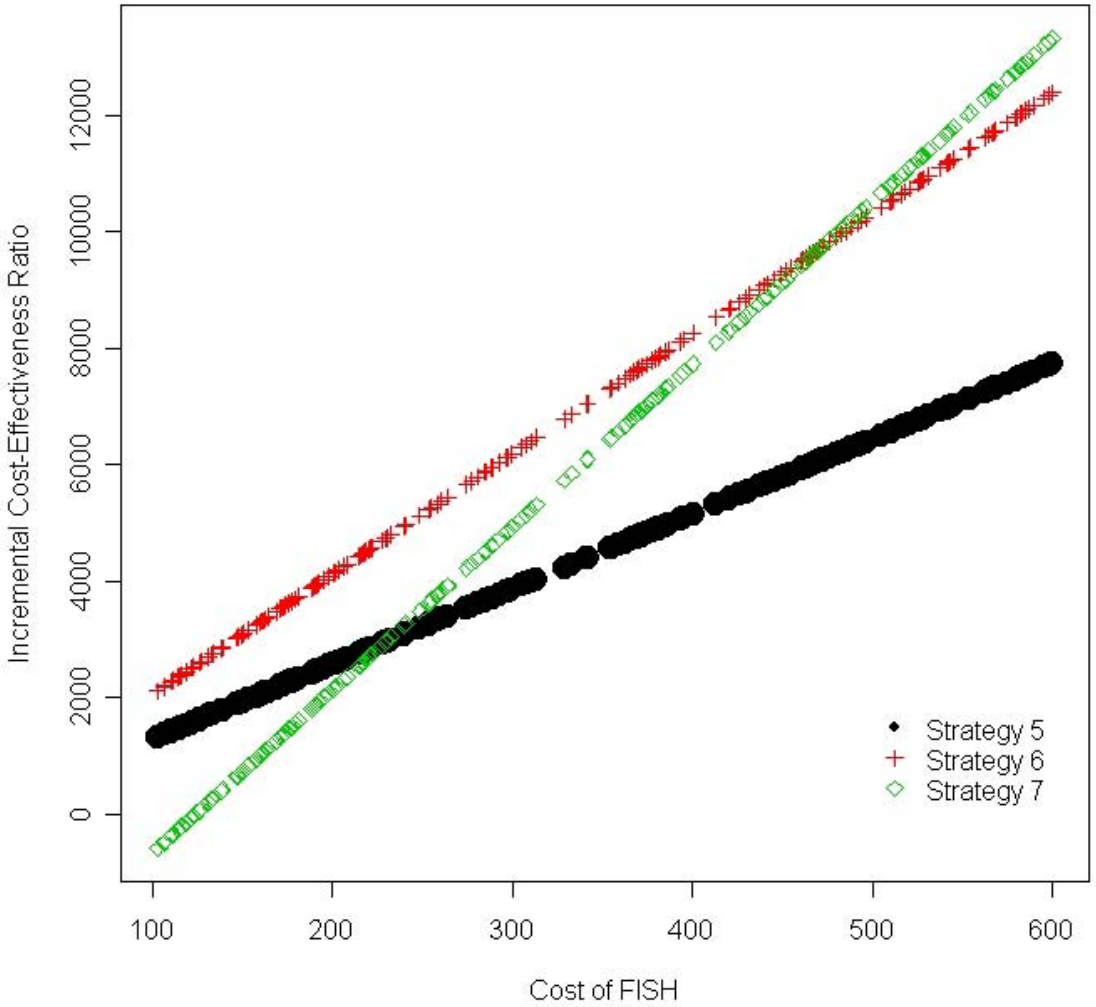


Figure 3d: Relation between ICER and cost of FISH*

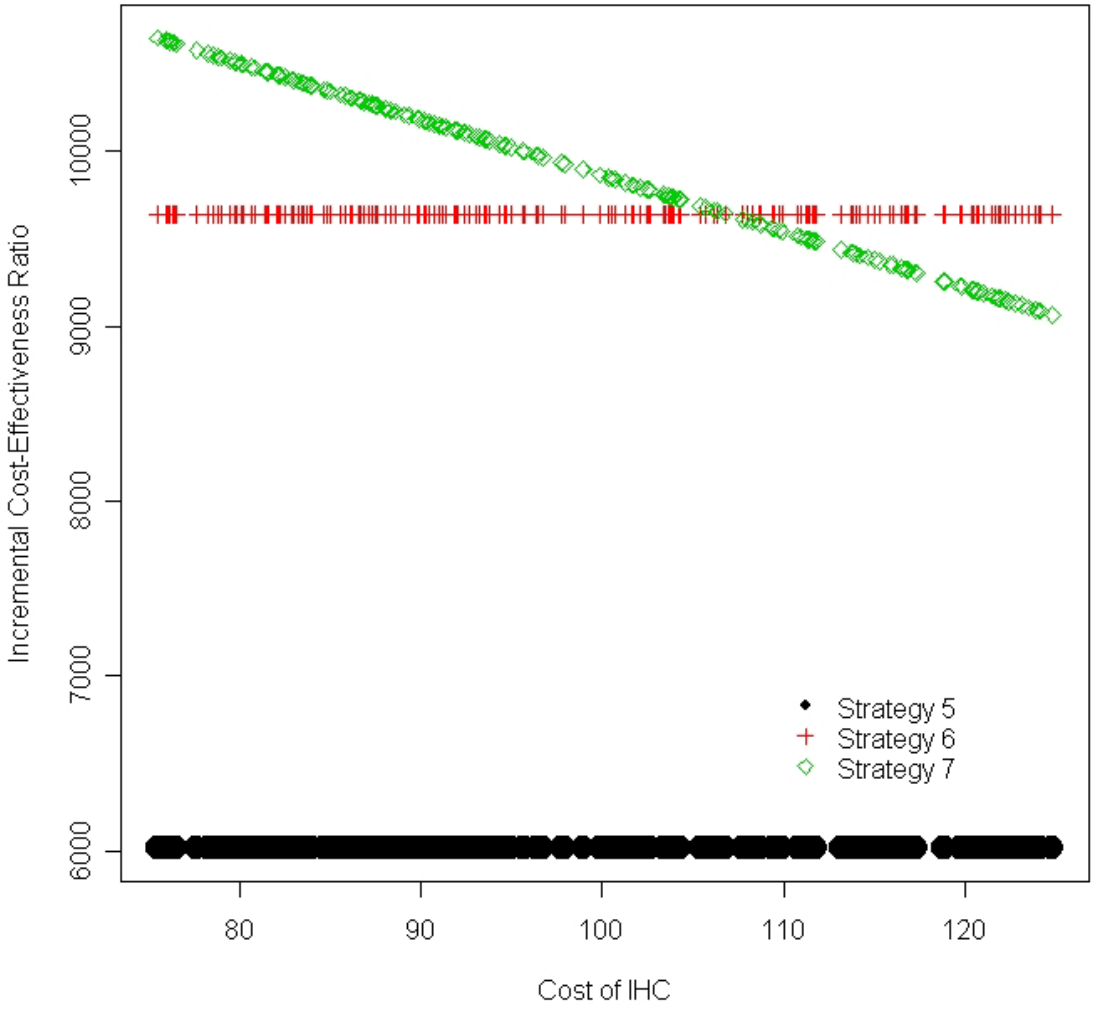
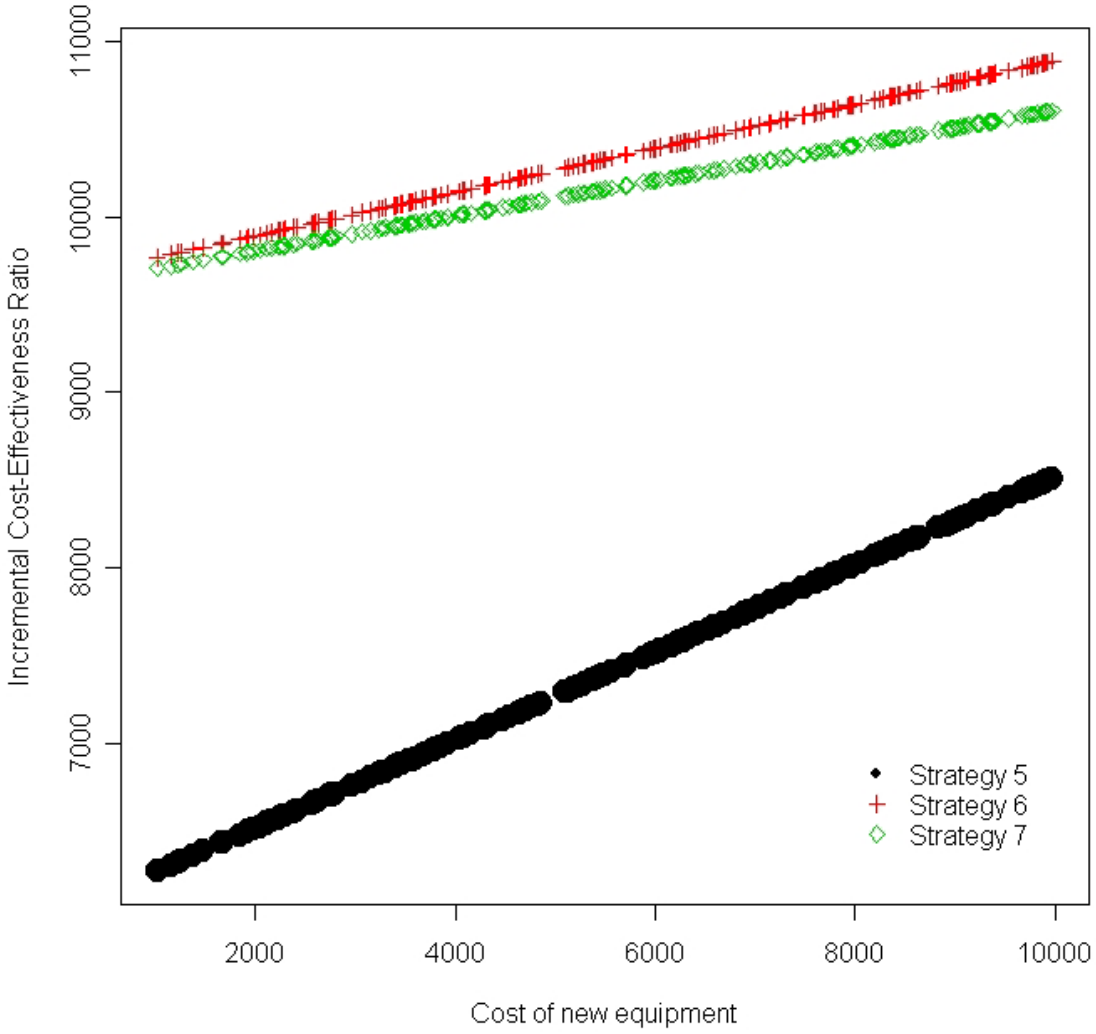


Figure 3e: Relation between ICER and annual cost of new equipment



DISCUSSION

The availability of an efficacious, yet expensive treatment for HER2 positive breast cancer underlines the need for accurate diagnosis of these patients. We reviewed the validity and reliability of the two most commonly used tests for HER2 detection: immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH). We also carried out a cost-effectiveness analysis that compared various testing strategies involving screening by IHC followed by confirmatory testing by FISH. Based on a meta-analysis of studies comparing IHC and FISH, we determined that the rate of false-positive IHC 3+ results could be as high as 8.1 (95% confidence interval: 4.1 %, 14.1%) translating into 4 women who would be exposed to the risks of Herceptin treatment with virtually no chance of gaining the benefits of the treatment.

Currently all breast cancer patients are tested for HER2 positive status with an IHC assay and those who receive a score of 2+ are confirmed using a FISH assay. Based on our review of the literature it appears that FISH is the superior test and we have assumed that FISH is the gold-standard for HER2 positive status. Thus, compared to the present policy the additional confirmatory testing of patients with other IHC results clearly increases the number of correctly diagnosed cases. In particular, the lowest ICER among such strategies results when IHC3+ cases are tested by FISH in addition to IHC2+. Importantly, this strategy will reduce the cost of Herceptin treatment (which we assumed to be effective and did not include in our analysis) by $200,000. The overall cost (median) of HER2 testing under this strategy would be $77,090 per year. Finally, our sensitivity analysis revealed that this strategy remains the most cost-effective unless the cost of FISH testing drops to less than $200/test. Adopting a more stringent strategy of FISH testing for all patients significantly increases the cost of testing but is expected to yield about 8 cases that would be missed by the current strategy. If all patients are to be routinely tested with FISH, the increased volume of testing will likely require new equipment.

Our conclusions are identical to those of Elkin et al[17], though their analysis was only among women in the metastatic phase of cancer and their analyses adjusted for the efficacy of trastuzumab. In a document prepared by the Programme de Gestion Therapeutique des Medicaments it was estimated that 100 patients would be treated with

trastuzumab at the MUHC. Our estimates, based on the literature, suggest that the number of women who are HER2 positive will be closer to 66 (95% CrI: 49, 85). The fact that the distribution of IHC scores among MUHC patients is comparable to our estimates based on a meta-analysis lends further support to our results. Further, our estimates of FISH positivity within categories of the IHC are similar to results obtained when evaluating an immunohistochemistry assay (CTA) that was developed specificially for the purpose of identifying women in three pivotal clinical trials carried out in women with meta-static breast cancer[42]. Thus one can assume that even if the standards of routinely used IHC tests were improved to match the CTA our conclusions would hold.

Though FISH has been demonstrated to have less than 100% sensitivity and specificity, it is widely regarded as the more valid test[29;32]. Retrospective analyses of data from clinical trials have demonstrated that the benefit of trastuzumab is limited to patients who are FISH positive[15]. Therefore we considered it a reasonable assumption to treat FISH as the gold-standard.

Besides cost-effectiveness, there are other reasons to be cautious about testing and treatment of women with IHC 1+ results with Herceptin. In particular, the ramifications of not treating IHC 0-1+ cases with trastuzumab is not known because such cases were not included in the HER2 trials. The available literature has shown that there is no significant difference in survival (without treatment) in IHC 0-1+ according to their FISH result[18]. The FISH positive cases in this group have generally been shown to have a low-level of amplification[18;39;40] So far there is no literature on the relation between trastuzumab efficacy and the level of amplification on the FISH test. Thus there is no proof of benefit for trastuzumab therapy for IHC 0,1+ and therefore hard to justify increasing the diagnostic costs for this group. Routine testing of IHC 0-1+ results would necessitate the purchase of new equipment given the substantial increase in the volume of testing. We found that purchase of equipment would increase the cost-effectiveness of a strategy of testing all patients with FISH over a strategy of confirmatory testing for patients with IHC scores of 1+, 2+ or 3+.

While our analysis has helped to clearly identify a more cost-effective testing strategy, a major limitation is that it is cross-sectional and does not account for the longitudinal costs and benefits of trastuzumab treatment. A more complete analysis

would account for the efficacy of trastuzumab and the increased life-expectancy (based on stage of cancer and the patient's age), the increased risk of cardiac toxicity with trastuzumab and the cost to patients who receive a false-negative result (IHC 0,1+) and are not treated. While we have assumed that FISH is a gold-standard a more complete analysis would take into account our knowledge of its sensitivity and specificity.

CONCLUSIONS

With the current arrangement for FISH testing external the MUHC, the most cost-effective strategy is to screen all patients with IHC, followed by confirmatory testing with FISH of those patients with IHC scores of 2+ or 3+. Purchase of new equipment may be justified if it is decided to test all patients with FISH.

RECOMMENDATION

If it is assumed that trastuzumab therapy will be offered, it is recommended that all breast cancer cases be screened with IHC and those who have scores of 2+ or 3+ be tested by FISH to confirm their HER2 positive status. In the unusual event that a breast cancer patient approaches the hospital with a positive FISH test carried out outside the MUHC, an IHC test needs to be carried out to confirm the course of action.

Appendix A: Summary of excluded studies

	First Author	Reason for exclusion			
		< 100 subjects	Non-licensed assays	Only subset tested with IHC and FISH/Selection bias	Did not follow manufacturer's instruction
1	Ainsworth[43]			X*	
2	Bankfalvi[44]	X	X		
3	Beatty[45]	X			
4	Bertucci[46]	X			
5	Bhargava[47]	X			
6	Birner[48]				X
7	Cianciulli[49]	X			
8	Courturier[50]		X		
9	Dandachi[51]			X	
10	Ellis[52]				X
11	Gancberg[53]				X
12	Ginestier[54]	X			
13	Gokhale[55]	X			
14	Hanna[56]			X	
15	Harris[57]		X		
16	Hauser-Kronberger[58]	X			
17	Jacobs[59]		X		
18	Jiminez[60]	X			
19	Lan[61]			X	
20	Lebeau[62]	X			
21	Lopez-Guerrero[63]	X			
22	Loring[34]	X			
23	Luftner[64]	X	X		

24	Nichols[65]	X			
25	Olsen[20]		X		
26	Pauletti[18]		X		
27	Perez[66]			X	
28	Ridolfi[67]		X		
29	Seidman[16]	X			
30	Thomson[68]	X			
31	Tse[12]	X			
32	Vincent-Saolomon[69]			X	
33	Wang[70]	X			
34	Wang[71]				X
35	Wixom[72]	X			
36	Zhao[73]	X			

* Only subjects with prospective follow-up were included. There were an unusually high number of FISH positive subjects in the IHC 1+ category.

Appendix B: Summary of included studies

Most studies used the PathVysion™ assay for FISH and the HercepTest™ assay for IHC. The study by Kakar et al used the Ventana Pathway™ assay for IHC.

1. Lottner et al, 2005[35]:

Selection process: Not clear how sample was selected

Concordance and/or reliability: Concordance only

Comments:

- The study also evaluated a multiparameter approach that allows simultaneous measurement of gene-amplification and protein overexpression. They conclude that this approach may be useful for predicting outcome of Herceptin treatment in patients with discordant FISH-IHC results.

2. Loring et al, 2005[34]:

Selection process: Consecutive breast tumor resection cases with recorded HercepTest™ scores on whole sections from August 2000 to April 2003.

Concordance and/or reliability: Concordance only

Comments:

- The study also evaluates the performance of the CISH assay for gene-amplification.
- They concluded that IHC 2+ and 3+ cases should be re-tested by a FISH or CISH assay.

3. Dowsett et al, 2003[30]:

Selection process: Samples received at laboratories specifically established to support trastuzumab treatment prior to its licensing in Europe.

Concordance and/or reliability: Both

Comments:

- In the reliability study all discrepancies involved the 2+ category.
- The study concluded that only IHC 2+ cases need to be re-tested with FISH as there are very few cases of discordance in the 0-1+ and 3+ cases. However, they caution

that their results hold only for large reference laboratories performing hundreds of IHC assays per year.

4. Press et al, 2002[32]:

Selection process: Specimens were selected if they had been previously molecularly characterized for HER2 gene amplification and over-expression using solid matrix blotting techniques, and if results of gene-amplification and over-expression were in agreement.

Concordance and/or reliability: Concordance

Comments:

- One of only two studies that compared commercially available FISH and IHC assays to a more accurate reference standard. They also compared these assays to two in-house IHC assays based on the 10H8 monoclonal antibody and the R60 polyclonal antibody.
- They concluded that FISH assays were significantly more sensitive and specific than the commercially available IHC assays, but were comparable to their in-house IHC assays.

5. Bartlett et al, 2002[29]:

Selection process: Breast carcinoma samples of patients with prospective follow-up were selected.

Concordance and/or reliability: Both

Comments:

- One of only two studies that compare commercial FISH and IHC assays to a more accurate standard. The reference standard used here was quantitative immunohistochemistry (Q-IHC).
- They conclude that FISH is a more accurate test than IHC. They recommend that FISH should be used both for selection of candidates for clinical trials as well as in routine practice.

6. Hoang et al, 2000[31]:

Selection process: 100 consecutive breast carcinoma for which the FISH test had been requested, and in which archival material was available.

Concordance and/or reliability: Both

Comments:

- Another IHC test using e2-4001 monoclonal antibody was also carried out.
- They also concluded that the HercepTest™ was the more reliable than the e2-4001 IHC assay and had better concordance with FISH.
- They concluded that the HercepTest™ was still a valuable screening tool given that most 0-1+ cases were FISH negative. However, they recommend that cases with IHC scores of 2+ and 3+ be re-tested with FISH due to poor reliability.

7. Kakar et al, 2000[38]:

Selection process: Unselected patients with invasive breast carcinoma treated and followed for a minimum of 50 months.

Concordance and /or reliability: Concordance

Comments:

- Only study included in this review that used the Pathway™ test for IHC.
- Recommend using IHC as a screening test and re-testing of IHC 2+ results with FISH.

8. Mrozkowiak et al, 2004[41]:

Selection process: Specimens obtained from patients operated for invasive breast cancer. Do not mention any other selection criteria. There were many more 2+ scores than in representative samples.

Concordance and/or reliability: Concordance only

Comments:

- Conclude that IHC 2+ cases alone should be retested by FISH at a recognized facility.

9. Yaziji et al, 2004[3]:

Selection process: Samples submitted for FISH testing between January 1999 and May 2003 were also tested with IHC.

Concordance and/or reliability: Concordance only

Comments:

- The failure rate for FISH was much higher than for IHC (5% vs 0.08%).
- Conclude that IHC should be used for screening patients for FISH testing.

10. Lal et al, 2004[39]:

Selection process: All tumors received for HER2 testing during a 20-month period including primary and metastatic breast carcinoma.

Concordance and/or reliability: Concordance only

Comments:

- They also evaluated the performance of a dual-color FISH assay.
- Found that almost 50% of FISH positive cases had low-level amplification and suggest that this may explain instances where there was no response to trastuzumab even in IHC3+ and FISH positive cases.
- Recommend re-testing only IHC 2+ results with FISH.

11. Ogura et al, 2004[37]:

Selection process: Consecutive patients with invasive ductal breast cancer who underwent surgical resection during an 11-month period.

Comments:

- Conclude that patients with IHC3+ score should be considered HER2 positive while those patients with an IHC2+ score should be re-tested with FISH.

12. Tsuda et al, 2002[36]:

Selection process: Consecutive patients who underwent mastectomy between 1992-1993 and 1999-2000 in Tokyo.

Concordance and/or reliability: Both

Comments:

- Found a very high reliability of FISH using various measures such as the HER2/CEP17 ratio and the HER2 copy number.
- The study concluded that IHC 2+ cases and low-level FISH amplification need to be interpreted carefully as these were the categories with the greatest discordance between IHC and FISH.

13. McCormick et al, 2003[40]:

Selection process: Samples received from diverse hospitals and clinics in the upper-midwest of the United States. There were roughly twice as many 2+ results (23%) than most studies. This may reflect a selection bias with 2+ cases being over-represented in this sample.

Concordance and/or reliability: Concordance

Comments:

- The study concluded that only IHC 2+ cases need to be re-tested with FISH.
- FISH+/IHC- cases mostly were typically low-copy amplified.

14. Roche et al, 2002[27]:

Selection process: A subset of 119 participants of the N9831 clinical trial of trastuzumab**.** Women were eligible for this trial if they had an IHC score of 0-2+ and a positive FISH score, or had an IHC score of 3+.

Concordance and/or reliability: Both

Comments:

- The authors were concerned by the high discordance between HercepTest results, that were used to screen patients for their trial, and the results of central IHC and FISH assays. Though they had too few FISH tests (only 9) there was an unacceptably high discordance in these results as well: 3/9. Following this study the authors altered their entry criteria into the trial requiring central confirmation of HER2 status.

15. Press et al, 2005[28]:

Selection process: Study subjects were participants in BCIRG 005, 006 and 007 trials.

Concordance and/or reliability: Both

Comments:

- Specimens obtained from women with operable breast cancer prior to metastasis (BCIRG 005 and 006) and previously un-treated women with metastatic breast cancer (BCIRG 007).
- The HercepTest was carried out by various labs while the FISH test was carried out at a central facility.
- Conclude that the better agreement between outside tests and central tests on FISH than HercepTest indicates that FISH is the better test. Recommend FISH testing for all women to identify candidates for trastuzumab treatment.

16. Dolan et al 2005[33]

Selection process: Biased sample of patients who were referred for FISH testing and who had IHC test results available.

Concordance and /or reliability: Concordance

Comments:

- Found a high discordance between FISH+ and IHC3+.
- Looked at the impact of specimen type and tumor grade on the discordance between IHC and FISH. The discordance was greater (79%) for core biopsy cases than in excisional biopsy cases (62%). The rates of discordance decreased with increasing tumor grade, from 85% for grade 1 (N=13), to 75% for grade 2 (N=40), and 53% for grade 3 (N=36).
- Recommend using FISH only if the number of IHC 2+ and 3+ cases is 1.6 to 2.6 times the number of IHC 0-1+ cases.

17. Paik et al, 2002[24]:

Selection process: A subset of 104 participants of the NSABP B-31 trial. Eligibility criteria required all women to have an IHC score of 3+.

Concordance and/or reliability: Reliability

Comments:

- There was a substantial disagreement between IHC results carried out at laboratories of accruing institutions and a central laboratory. In most cases (80/104) the IHC was

carried out using the Herceptest assay. A total of 22 of the 104 samples (21.2%) could not be confirmed by one of either Herceptest or a PathVysion FISH assay at the central laboratory.

- The authors concluded that there was a strong association between reliability and volume of testing. Eighteen of 75 (24%) of tests from smaller volume laboratories (<100 cases per month) could not be confirmed compared to 1 of 29 (3%) from larger volume laboratories.

18. Dybdal et al 2005[42]:

Selection process: Patients were participants in three pivotal clinical trials for trastuzumab among women with metastatic breast cancer. The sample used for this study, however, was not random but limited to cases where sufficient tissue was available for a second confirmatory series of FISH assays.

Concordance and/or reliability: Reliability of FISH

Comments:

- A large percentage of cases were non-informative. In one laboratory where non-informative cases were not re-tested there was a 13% failure rate. While in the second laboratory, where FISH was repeated for non-informative assays following individual optimization of FISH conditions, the failure rate was 8%.
- They also evaluated the concordance between the FISH assay and the clinical trials IHC assay (CTA) and concluded there was good concordance between the two.
- Recommend that FISH be used for selecting women for trastuzumab therapy.

REFERENCES

1. Hicks DG,.Tubbs RR. Assessment of the HER2 status in breast cancer by fluorescence in situ hybridization: a technical review with interpretive guidelines. *Hum Pathol* 2005;36:250-61.

2. Ross JS,.Fletcher JA. The HER-2/neu oncogene in breast cancer: prognostic factor, predictive factor, and target for therapy. *Stem Cells* 1998;16:413-28.

3. Yaziji H, Goldstein LC, Barry TS, Werling R, Hwang H, Ellis GK *et al*. HER-2 testing in breast cancer using parallel tissue-based methods. *JAMA* 2004;291:1972-7.

4. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A *et al*. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 2001;344:783-92.

5. Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L *et al*. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J Clin Oncol* 2002;20:719-26.

6. Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, Fehrenbacher L *et al*. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J Clin Oncol* 1999;17:2639-48.

7. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE Jr, Davidson NE *et al*. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005;353:1673-84.

8. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I *et al*. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 2005;353:1659-72.
   Notes: CORPORATE NAME: Herceptin Adjuvant (HERA) Trial Study Team.

9. Programme de Gestion Thérapeutique des Médicaments (PGTM). Prise de Position Préliminaire Trastuzumab (HERCEPTIN) dans le Traitement adjusvant du Cancer du Sein. 2005.

10. Comité de l'évolution des pratiques en oncologie (CEPO). Guide dutilisation du trastuzumab (Herceptin) dans le traitement ajduvant du cancer du sein. 2005.

11. Barron, Hal. Genentech biooncology Important Drug Warning. 2005.

12. Dressler LG, Berry DA, Broadwater G, Cowan D, Cox K, Griffin S *et al*. Comparison of HER2 status by fluorescence in situ hybridization and immunohistochemistry to predict benefit from dose escalation of adjuvant doxorubicin-based therapy in node-positive breast cancer patients. *J Clin Oncol* 2005;23:4287-97.

13. Bartlett J, Mallon E, Cooke T. The clinical evaluation of HER-2 status: which test to use? *J Pathol* 2003;199:411-7.

14. Bilous M, Dowsett M, Hanna W, Isola J, Lebeau A, Moreno A *et al*. Current perspectives on HER2 testing: a review of national testing guidelines. *Mod Pathol* 2003;16:173-82.

15. Mass RD, Press MF, Anderson S, Cobleigh MA, Vogel CL, Dybdal N *et al*. Evaluation of clinical outcomes according to HER2 detection by fluorescence in situ hybridization in women with metastatic breast cancer treated with trastuzumab. *Clin Breast Cancer* 2005;6:240-6.

16. Seidman AD, Fornier MN, Esteva FJ, Tan L, Kaptain S, Bach A *et al*. Weekly trastuzumab and paclitaxel therapy for metastatic breast cancer with analysis of efficacy by HER2 immunophenotype and gene amplification. *J Clin Oncol* 2001;19:2587-95.

17. Elkin EB, Weinstein MC, Winer EP, Kuntz KM, Schnitt SJ, Weeks JC. HER-2 testing and trastuzumab therapy for metastatic breast cancer: a cost-effectiveness analysis. *J Clin Oncol* 2004;22:854-63.

18. Pauletti G, Dandekar S, Rong H, Ramos L, Peng H, Seshadri R *et al*. Assessment of methods for tissue-based detection of the HER-2/neu alteration in human breast cancer: a direct comparison of fluorescence in situ hybridization and immunohistochemistry. *J Clin Oncol* 2000;18:3651-64.

19. Dranitsaris G., Norris B., Hanna W., O'Malley F., Gelmon K. Identifying the optimal timing of HER2 testing in patients with breast cancer: a Canadian economic evaluation. *Current Oncology*10:36.

20. Norum J, Risberg T, Olsen JA. A monoclonal antibody against HER-2 (trastuzumab) for metastatic breast cancer: a model-based cost-effectiveness analysis. *Ann Oncol* 2005;16:909-14.

21. Neyt MJ, Albrecht JA, Clarysse B, Cocquyt VF. Cost-effectiveness of Herceptin: a standard cost model for breast-cancer treatment in a Belgian university hospital. *Int J Technol Assess Health Care* 2005;21:132-7.

22. Lewis, R., Bagnall, A-M, Forbes, C, Shirran, E, Duffy, S, Kleijnen, J, ter Riet, G, and Riemsma, R. A rapid and systematic review of the clinical effectiveness and cost-

effectiveness of trastuzumab for breast cancer . 2002. National Institute for Clinical Excellence.

23. National Institutes for Clinical Excellence. Guidance on the use of trastuzumab for the treatment of advanced breast cancer. 2002.

24. Paik S, Bryant J, Tan-Chiu E, Romond E, Hiller W, Park K *et al*. Real-world performance of HER2 testing--National Surgical Adjuvant Breast and Bowel Project experience. *J Natl Cancer Inst* 2002;94:852-4.

25. Cohen J. A coefficient of agreement for nominal scales . *Educational and Psychological Measurement*
1960;20:37-46.

26. Begg CB,.Greenes Robert A. Assessment of Diagnostic Tests When Disease Verification is Subject to Selection Bias. *Biometrics* 1983;39:207-15.

27. Roche PC, Suman VJ, Jenkins RB, Davidson NE, Martino S, Kaufman PA *et al*. Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. *J Natl Cancer Inst* 2002;94:855-7.

28. Press MF, Sauter G, Bernstein L, Villalobos IE, Mirlacher M, Zhou JY *et al*. Diagnostic evaluation of HER-2 as a molecular target: an assessment of accuracy and reproducibility of laboratory testing in large, prospective, randomized clinical trials. *Clin Cancer Res* 2005;11:6598-607.

29. Bartlett JM, Going JJ, Mallon EA, Watters AD, Reeves JR, Stanton P *et al*. Evaluating HER2 amplification and overexpression in breast cancer. *J Pathol* 2001;195:422-8.

30. Dowsett M, Bartlett J, Ellis IO, Salter J, Hills M, Mallon E *et al*. Correlation between immunohistochemistry (HercepTest) and fluorescence in situ hybridization (FISH) for HER-2 in 426 breast carcinomas from 37 centres. *J Pathol* 2003;199:418-23.

31. Hoang MP, Sahin AA, Ordonez NG, Sneige N. HER-2/neu gene amplification compared with HER-2/neu protein overexpression and interobserver reproducibility in invasive breast carcinoma. *Am J Clin Pathol* 2000;113:852-9.

32. Press MF, Slamon DJ, Flom KJ, Park J, Zhou JY, Bernstein L. Evaluation of HER-2/neu gene amplification and overexpression: comparison of frequently used assay methods in a molecularly characterized cohort of breast cancer specimens. *J Clin Oncol* 2002;20:3095-105.

33. Dolan M,.Snover D. Comparison of immunohistochemical and fluorescence in situ

hybridization assessment of HER-2 status in routine practice. *Am J Clin Pathol* 2005;123:766-70.

34. Loring P, Cummins R, O'Grady A, Kay EW. HER2 positivity in breast carcinoma: a comparison of chromogenic in situ hybridization with fluorescence in situ hybridization in tissue microarrays, with targeted evaluation of intratumoral heterogeneity by in situ hybridization. *Appl Immunohistochem Mol Morphol* 2005;13:194-200.

35. Lottner C, Schwarz S, Diermeier S, Hartmann A, Knuechel R, Hofstaedter F *et al*. Simultaneous detection of HER2/neu gene amplification and protein overexpression in paraffin-embedded breast cancer. *J Pathol* 2005;205:577-84.

36. Tsuda H, Akiyama F, Terasaki H, Hasegawa T, Kurosumi M, Shimadzu M *et al*. Detection of HER-2/neu (c-erb B-2) DNA amplification in primary breast carcinoma. Interobserver reproducibility and correlation with immunohistochemical HER-2 overexpression. *Cancer* 2001;92:2965-74.

37. Ogura H, Akiyama F, Kasumi F, Kazui T, Sakamoto G. Evaluation of HER-2 status in breast carcinoma by fluorescence in situ hybridization and immunohistochemistry. *Breast Cancer* 2003;10:234-40.

38. Kakar S, Puangsuvan N, Stevens JM, Serenas R, Mangan G, Sahai S *et al*. HER-2/neu assessment in breast cancer by immunohistochemistry and fluorescence in situ hybridization: comparison of results and correlation with survival. *Mol Diagn* 2000;5:199-207.

39. Lal P, Salazar PA, Hudis CA, Ladanyi M, Chen B. HER-2 testing in breast cancer using immunohistochemical analysis and fluorescence in situ hybridization: a single-institution experience of 2,279 cases and comparison of dual-color and single-color scoring. *Am J Clin Pathol* 2004;121:631-6.

40. McCormick SR, Lillemoe TJ, Beneke J, Schrauth J, Reinartz J. HER2 assessment by immunohistochemical analysis and fluorescence in situ hybridization: comparison of HercepTest and PathVysion commercial assays. *Am J Clin Pathol* 2002;117:935-43.

41. Mrozkowiak A, Olszewski WP, Piascik A, Olszewski WT. HER2 status in breast cancer determined by IHC and FISH: comparison of the results. *Pol J Pathol* 2004;55:165-71.

42. Dybdal N, Leiberman G, Anderson S, McCune B, Bajamonde A, Cohen RL *et al*. Determination of HER2 Gene Amplification by Fluorescence In situ Hybridization and Concordance with the Clinical Trials Immunohistochemical Assay in Women with Metastatic Breast Cancer Evaluated for Treatment with Trastuzumab. *Breast Cancer Res Treat* 2005;93:3-11.

43. Ainsworth R, Bartlett JM, Going JJ, Mallon EA, Forsyth A, Richmond J *et al*. IHC for Her2 with CBE356 antibody is a more accurate predictor of Her2 gene amplification by FISH than HercepTest in breast carcinoma. *J Clin Pathol* 2005;58:1086-90.

44. Bankfalvi A, Boecker W, Reiner A. Comparison of automated and manual determination of HER2 status in breast cancer for diagnostic use: a comparative methodological study using the Ventana BenchMark automated staining system and manual tests. *Int J Oncol* 2004;25:929-35.

45. Beatty BG, Bryant R, Wang W, Ashikaga T, Gibson PC, Leiman G *et al*. HER-2/neu detection in fine-needle aspirates of breast cancer: fluorescence in situ hybridization and immunocytochemical analysis. *Am J Clin Pathol* 2004;122:246-55.

46. Bertucci F, Borie N, Ginestier C, Groulet A, Charafe-Jauffret E, Adelaide J *et al*. Identification and validation of an ERBB2 gene expression signature in breast cancers. *Oncogene* 2004;23:2564-75.

47. Bhargava R, Naeem R, Marconi S, Luszcz J, Garb J, Gasparini R *et al*. Tyrosine kinase activation in breast carcinoma with correlation to HER-2/neu gene amplification and receptor overexpression. *Hum Pathol* 2001;32:1344-50.

48. Birner P, Oberhuber G, Stani J, Reithofer C, Samonigg H, Hausmaninger H *et al*. Evaluation of the United States Food and Drug Administration-approved scoring and test system of HER-2 protein expression in breast cancer. *Clin Cancer Res* 2001;7:1669-75. Notes: CORPORATE NAME: Austrian Breast & Colorectal Cancer Study Group.

49. Cianciulli AM, Botti C, Coletta AM, Buglioni S, Marzano R, Benevolo M *et al*. Contribution of fluorescence in situ hybridization to immunohistochemistry for the evaluation of HER-2 in breast cancer. *Cancer Genet Cytogenet* 2002;133:66-71.

50. Couturier J, Vincent-Salomon A, Nicolas A, Beuzeboc P, Mouret E, Zafrani B *et al*. Strong correlation between results of fluorescent in situ hybridization and immunohistochemistry for the assessment of the ERBB2 (HER-2/neu) gene status in breast carcinoma. *Mod Pathol* 2000;13:1238-43.

51. Dandachi N, Dietze O, Hauser-Kronberger C. Evaluation of the clinical significance of HER2 amplification by chromogenic in situ hybridisation in patients with primary breast cancer. *Anticancer Res* 2004;24:2401-6.

52. Ellis CM, Dyson MJ, Stephenson TJ, Maltby EL. HER2 amplification status in breast cancer: a comparison between immunohistochemical staining and fluorescence in situ hybridisation using manual and automated quantitative image analysis scoring techniques. *J Clin Pathol* 2005;58:710-4.

53. Gancberg D, Jarvinen T, di Leo A, Rouas G, Cardoso F, Paesmans M *et al*. Evaluation of HER-2/NEU protein expression in breast cancer by immunohistochemistry: an interlaboratory study assessing the reproducibility of HER-2/NEU testing. *Breast Cancer Res Treat* 2002;74:113-20.

54. Ginestier C, Charafe-Jauffret E, Penault-Llorca F, Geneix J, Adelaide J, Chaffanet M *et al*. Comparative multi-methodological measurement of ERBB2 status in breast cancer. *J Pathol* 2004;202:286-98.

55. Gokhale S, Gatalica Z, Mohammad A, Rampy AI, Velagaleti Gopalrao VN. FISH for HER-2/neu in breast cancer: standardization makes the difference! *Indian J Cancer* 2004;41:152-8.

56. Hanna WM,.Kwok K. Chromogenic in-situ hybridization: a viable alternative to fluorescence in-situ hybridization in the HER2 testing algorithm. *Mod Pathol* 2006;19:481-7.

57. Harris LN, Liotcheva V, Broadwater G, Ramirez MJ, Maimonis P, Anderson S *et al*. Comparison of methods of measuring HER-2 in metastatic breast cancer patients treated with high-dose chemotherapy. *J Clin Oncol* 2001;19:1698-706.

58. Hauser-Kronberger C,.Dandachi N. Comparison of chromogenic in situ hybridization with other methodologies for HER2 status assessment in breast cancer. *J Mol Histol* 2004;35:647-53.

59. Jacobs TW, Gown AM, Yaziji H, Barnes MJ, Schnitt SJ. Comparison of fluorescence in situ hybridization and immunohistochemistry for the evaluation of HER-2/neu in breast cancer. *J Clin Oncol* 1999;17:1974-82.

60. Jimenez RE, Wallis T, Tabasczka P, Visscher DW. Determination of Her-2/Neu status in breast carcinoma: comparative analysis of immunohistochemistry and fluorescent in situ hybridization. *Mod Pathol* 2000;13:37-45.

61. Lan C, Liu JM, Liu TW, Hsu DH, Liang S, Chen JR *et al*. erb-b2 amplification by fluorescence in situ hybridization in breast cancer specimens read as 2+ in immunohistochemical analysis. *Am J Clin Pathol* 2005;124:97-102.

62. Lebeau A, Deimling D, Kaltz C, Sendelhofert A, Iff A, Luthardt B *et al*. Her-2/neu analysis in archival tissue samples of human breast cancer: comparison of immunohistochemistry and fluorescence in situ hybridization. *J Clin Oncol* 2001;19:354-63.

63. Lopez-Guerrero JA, Navarro S, Noguera R, Almenar S, Pellin A, Vazquez C *et al*. Histological tumor grade correlates with HER2/c-erB-2 status in invasive breast cancer: a

comparative analysis between immunohistochemical (CB11 clone and Herceptest), FISH and differential PCR procedures. *Arkh Patol* 2003;65:50-5.

64. Luftner D, Henschke P, Kafka A, Anagnostopoulos I, Wiechen K, Geppert R *et al*. Discordant results obtained for different methods of HER-2/neu testing in breast cancer--a question of standardization, automation and timing. *Int J Biol Markers* 2004;19:1-13.

65. Nichols DW, Wolff DJ, Self S, Metcalf JS, Jacobs D, Kneuper-Hall R *et al*. A testing algorithm for determination of HER2 status in patients with breast cancer. *Ann Clin Lab Sci* 2002;32:3-11.

66. Perez EA, Roche PC, Jenkins RB, Reynolds CA, Halling KC, Ingle JN *et al*. HER2 testing in patients with breast cancer: poor correlation between weak positivity by immunohistochemistry and gene amplification by fluorescence in situ hybridization. *Mayo Clin Proc* 2002;77:148-54.

67. Ridolfi RL, Jamehdor MR, Arber JM. HER-2/neu testing in breast carcinoma: a combined immunohistochemical and fluorescence in situ hybridization approach. *Mod Pathol* 2000;13:866-73.

68. Thomson TA, Hayes MM, Spinelli JJ, Hilland E, Sawrenko C, Phillips D *et al*. HER-2/neu in breast cancer: interobserver variability and performance of immunohistochemistry with 4 antibodies compared with fluorescent in situ hybridization. *Mod Pathol* 2001;14:1079-86.

69. Vincent-Salomon A, MacGrogan G, Couturier J, Arnould L, Denoux Y, Fiche M *et al*. Calibration of immunohistochemistry for assessment of HER2 in breast cancer: results of the French multicentre GEFPICS study. *Histopathology* 2003;42:337-47.

70. Wang S, Saboorian MH, Frenkel E, Hynan L, Gokaslan ST, Ashfaq R. Laboratory assessment of the status of Her-2/neu protein and oncogene in breast cancer specimens: comparison of immunohistochemistry assay with fluorescence in situ hybridisation assays. *J Clin Pathol* 2000;53:374-81.

71. Wang S, Saboorian MH, Frenkel EP, Haley BB, Siddiqui MT, Gokaslan S *et al*. Assessment of HER-2/neu status in breast cancer. Automated Cellular Imaging System (ACIS)-assisted quantitation of immunohistochemical assay achieves high accuracy in comparison with fluorescence in situ hybridization assay as the standard. *Am J Clin Pathol* 2001;116:495-503.

72. Wixom CR, Albers EA, Weidner N. Her2 amplification: correlation of chromogenic in situ hybridization with immunohistochemistry and fluorescence in situ hybridization. *Appl Immunohistochem Mol Morphol* 2004;12:248-51.

73. Zhao J, Wu R, Au A, Marquez A, Yu Y, Shi Z. Determination of HER2 gene amplification by chromogenic in situ hybridization (CISH) in archival breast carcinoma. *Mod Pathol* 2002;15:657-65.