

sBLA
STN: 103792\5008\0
[Original BLA 98-0369]

Trastuzumab
Herceptin®
(rhuMAbHER2)

Date of Submission: April 2, 2001
Final Approval: Decision Pending

Sponsor: Genentech, Inc.

Food and Drug Administration
Center for Biologics Evaluation and Research

FDA/CBER Review Team

Clinical Reviewer: Susan Jerian, MD

Statistical Reviewer: Chao Wang, PhD

BIMO Reviewer: Deborah Bower

Table of Contents

I.	Introduction	4
II.	Background	5
III.	FISH Studies Conducted	6
IV.	Results of FISH Testing	12
V.	Conclusions	33
VI.	Appendix A: Exploratory Analyses	35
VII.	Appendix B: PathVysion Package Insert	45
VIII.	Appendix C: HercepTest Package Insert	46
IX.	Appendix D: Trastuzumab Package Insert	47

I. INTRODUCTION

The sponsor of this supplemental biologic license application (sBLA), Genentech, Inc., requests that information on use of an assay to detect gene amplification, fluorescence in situ hybridization (FISH), be added to the trastuzumab (Herceptin) package insert (PI). The original licensing clinical trials upon which the approval of trastuzumab was based, enrolled patients with HER2 protein overexpression as determined by an immunohistochemistry assay (IHC). The trastuzumab approval was necessarily linked to the approval of the diagnostic immunohistochemistry (IHC) assay. FDA/CBER required that a diagnostic be available to clinicians (commercially or otherwise) for instances where the decision to treat a patient with a therapeutic product was dependent upon using the diagnostic assay.

The approved assay for the detection of HER2 protein overexpression, HercepTest manufactured by DAKO, was licensed at the same time as trastuzumab. HercepTest was approved based upon a concordance study comparing HercepTest to the Clinical Trial Assay (CTA). The CTA was the immunohistochemistry assay used to determine patient eligibility for the trastuzumab clinical trials; it employed two monoclonal antibodies: 4D5, the parent antibody to trastuzumab, and CB11. HercepTest employed a polyclonal antibody. IHC scores of 2+ and 3+ (on a scale of 0, 1+, 2+, and 3+) were considered positive for overexpression in both assays. HercepTest was not tested on tissue samples from the clinical trial, because the only archived tissue samples available were pre-cut slides upon which conduct of IHC was felt to be technically unreliable.

At the 1998 Oncologic Drugs Advisory Committee Meeting which reviewed trastuzumab, exploratory analyses of the clinical trial data were presented and strongly suggested that the clinical benefit of trastuzumab was primarily seen in patients whose tumors were scored as IHC 3+ using the CTA. While the committee members noted this as an important finding, they also noted that there may be a subset of 2+ patients for whom therapy might be effective and they did not want the PI to exclude 2+ patients from consideration for trastuzumab therapy. FDA followed this advice by not restricting the indication to 3+ patients and included the data for these subgroups (2+ and 3+) in the trastuzumab package insert to assist clinicians in treatment decisions.

Facing many unanswered questions regarding HER2 detection systems and with the anticipation of further developments in the field, FDA asked Genentech to make the following postmarketing commitment in 1998:

“To assess the clinical outcome of patients selected for treatment on the basis of the DAKO test [HercepTest] and other HER2 diagnostics in the context of Herceptin clinical trials.”

While this supplemental biologic license application (sBLA) application does not fulfill this commitment in total, FDA recognizes the extensive use of FISH testing off

label for the selection of patients to be treated with trastuzumab and feels it is in the interest of the public health to comment on the use of FISH in the trastuzumab label. FDA seeks the advice of the ODAC at this time regarding the type of information to include in the trastuzumab PI and what additional trial designs might provide useful data to enable more definitive recommendations to clinicians about HER testing.

II. BACKGROUND

A. Regulatory Timeline for the FISH sBLA

In March 2000, Genentech informed CBER that they had conducted exploratory, retrospective concordance and clinical outcomes studies using fluorescence in situ hybridization (FISH) as a method for selecting HER2 amplified patients to be treated with trastuzumab. These studies were undertaken by Genentech, and not initially by the FISH assay manufacturer. They were not conducted under IND. The sponsor felt the results to be sufficiently provocative to warrant a proposal for a sBLA to allow FISH to be used to select patients for trastuzumab therapy. In light of developments in FISH assays and more extensive off label use of FISH to select patients for trastuzumab therapy, FDA agreed to review the initial proposal. Upon review it was found that the proposal was not acceptable, because there was a significant amount of missing data in the clinical outcome study and this missing data appeared to occur in a non-random fashion. FDA also advised the sponsor that in order to make predictive or comparative claims in the PI, a prospective trial would need to be conducted.

In August 2000, FDA agreed with Genentech's proposal to try to minimize the amount of missing data by performing FISH on previously stained slides based on methods developed by Dr. Michael Press at the University of Southern California (USC). FDA advised the sponsor that even with this data, the clinical outcome data may at best be supportive and may not rise to the level of robustness to be included in the package insert. The sBLA proposing the use of FISH (PathVysion assay kit) as an aid in the selection of patients for trastuzumab therapy was received in April 2001. At approximately the same time, a Pre-Market Approval application (PMA) was also filed in the Center for Devices and Radiological Health (CDRH) by Vysis, the manufacturer of PathVysion.

B. HER2 Amplification as Detected by FISH

There are currently two approved FISH assay kits (PathVysion and INFORM) for the detection of HER2 amplification, but neither is indicated as an aid in the selection of patients for treatment with trastuzumab. Genentech employed the PathVysion FISH assay in their studies. PathVysion by Vysis is presently indicated "as an adjunct to existing clinical and pathologic information currently used as prognostic factors in stage II, node-positive breast cancer patients" and "as an aid to predict disease-free and overall survival in patients with stage II, node positive breast cancer treated with

adjuvant cyclophosphamide, doxorubicin, and 5-fluorouracil (CAF) chemotherapy.”
See Appendix B for PathVysion package insert.

C. Nature of the Data

FDA makes the following observations about the nature of the data submitted in this application. They are exploratory, retrospective data with provocative results and are, in part, appropriate for inclusion in the trastuzumab label for the purpose of supplying clinicians with information upon which they may base their decisions for individual patient treatment or design of clinical trials. The data are not from prospective, randomized, double-blinded, controlled, multi-center trials and can neither answer questions regarding the predictive capability of FISH to select patients who would most benefit from trastuzumab therapy, nor answer questions regarding a hypothesized comparative benefit of FISH vs IHC. These data do not address the issue of the clinical utility of administering trastuzumab therapy to patients whose tumors test as FISH positive and IHC negative. As such, they do not rise to the level for expansion of the indications in the trastuzumab label.

It is also the perception of the FDA that there is considerable confusion and misunderstanding on the part of the oncology community regarding the conduct, performance characteristics, clinical application and interpretation of IHC and FISH. These may be significant enough to warrant general precautionary comment in the trastuzumab label for the sake of clarification of these issues.

FDA views both IHC and FISH as semi-quantitative if performed under ideal circumstances. Both methods require subjective interpretation of special stains.

III. FISH STUDIES CONDUCTED

A. Overview

The FISH studies conducted on the archived clinical trial tissue samples proceeded in a progressive and complex fashion. The sponsor conducted retrospective concordance, clinical outcome, and validation studies. The first study undertaken was the concordance study comparing detection of gene amplification to protein overexpression. Gene amplification was assessed by LabCorp using a FISH assay kit (PathVysion). Protein overexpression was previously assessed by LabCorp using the CTA. Based upon the results of the concordance study, the clinical outcome study was initiated on tissues samples from patients for whom there were at least 2 unstained slides in storage. While the results of the clinical outcome study were of interest, there was a considerable amount of missing data due to the lack of sufficient number of unstained slides available and failure of the FISH assay in some cases where unstained slides were available. Efforts to fill in the missing data were made

by employing the services of Dr. Michael Press who had experience with performing FISH on previously stained slides. The clinical outcome study was then expanded to include assessments at the lab of Dr. Press. In conjunction with this testing was a validation study to compare the results from the LabCorp testing to those with the Press lab testing. The validation study revealed potential systemic differences between the labs leading to further validation testing between LabCorp and Press. Below are descriptions of the various studies followed by flow diagrams outlining where tissue samples were assessed.

FISH assay used on all samples: PathVysion manufactured by Vysis. See package insert in Appendix B .

Laboratory sites:

Commercial Laboratory site = Laboratory Corporation (LabCorp)
Academic Laboratory site = Dr. Michael Press's laboratory at the
University of Southern California (Press)

Retrospective studies conducted:

CTA vs FISH Concordance Study:	samples from all screened patients, (IHC scores of 0, 1+, 2+, 3+)
Clinical Outcome Study:	samples from H0648g, H0649g, and H0650g (2+ and 3+ only)
LabCorp vs Press Validation:	samples from H0648g, H0649g, and H0650g (2+ and 3+ only)

Tissue source: Archived pre-cut slides from previously conducted trastuzumab licensing clinical trials H0648g and H0649g, non-licensing Phase 2 trial H650g, and samples used to screen patients for trastuzumab clinical development program prior to approval were used. Slides may have been cut and unstained or cut and previously stained as IgG control slides. Slides used were from patients' tissue samples previously tested by LabCorp using CTA (IHC).

Description of clinical trials from which tissue samples were obtained:

H0648g: Chemotherapy and Antibody Response Evaluation (CARE): A Phase III, Multinational, Randomized Study of Recombinant Humanized Anti-p185^{HER2} Monoclonal Antibody (rhuMAb HER2) Combined with Chemotherapy in Patients with HER2 Overexpression who have not Received Cytotoxic Chemotherapy for Metastatic Breast Cancer.

Enrollment conducted from June 12, 1995 to March 7, 1997; patient follow up through October 1999.

Multicenter, randomized, open label trial comparing chemotherapy in combination with trastuzumab vs chemotherapy alone. Depending on prior anthracycline use in the adjuvant setting, patients were either treated with standard AC (anthracycline plus cyclophosphamide) therapy if no prior anthracycline or Paclitaxel as the chemotherapy portion of the therapy.

Enrolled population: Patients with metastatic breast cancer who had not received prior chemotherapy for their metastatic disease and whose tumor tissue tested as 2+ or 3+ by immunohistochemistry assay using CTA.

The primary endpoint was median time to progression as determined by independent response evaluation committee with secondary endpoints of overall response rate, median duration of response, time to treatment failure, overall survival and quality of life.

Results are as noted in the current package insert (Appendix D).

H0649g: A Multinational, Open-Label Study of Recombinant Humanized Anti-p185^{HER2} Monoclonal Antibody (rhuMAb HER2) in Patients with HER2/neu Overexpression Who Have Relapsed Following One or Two Cytotoxic Chemotherapy Regimens for Metastatic Breast Cancer

Conducted April 24, 1995 to June 4, 1997

Multicenter, single arm trial of trastuzumab as single agent therapy.

Enrolled population: Patients with metastatic disease who had received treatment with one or two prior chemotherapy regimens and whose tumor tissue tested as 2+ or 3+ by immunohistochemistry assay using CTA.

The primary endpoint was overall objective response rate (complete response plus partial response) as determined by independent response evaluation committee with secondary endpoints of median duration of response, median time to progression, median time to treatment failure, overall survival and quality of life.

Results are as noted in the current package insert (Appendix D).

H0650g: A Multinational, Randomized, Single-Blind Study of Recombinant Humanized Anti-p198HER2 Monoclonal Antibody (rhuMAb HER2) in Patients with HER2/neu Overexpression Who Have Not Received Prior Cytotoxic Chemotherapy for Metastatic Breast Cancer

This trial enrolled patients who were not eligible to receive chemotherapy or who electively chose not to receive chemotherapy as front line therapy for their metastatic disease. All patients received trastuzumab as a single agent; randomization was to one of two dose levels.

Clinical efficacy data from this trial was not previously accepted or reviewed by the FDA for inclusion in the package insert because the trial was conducted in a subset of primary metastatic breast cancer patients from which conclusions about use of trastuzumab as single agent first line therapy could not be ascertained. Study H0659g data for this FISH sBLA were reviewed by the FDA for concordance with CTA and for comparisons of the two testing facilities, but not for clinical outcome.

B. CTA vs FISH Concordance Study

LabCorp Protocol: Retrospective Study

Primary Endpoint: Determine the level of concordance between two methods of measuring HER2 overexpression in patients with metastatic breast cancer: 1) Clinical Trial Assay (CTA) immunohistochemistry (IHC) utilized in the trastuzumab clinical trials and 2) Fluorescence *in situ* Hybridization (FISH). Concordance of > 75% acceptable.

Study Design: Retrospective, single center, single blind study of 600 randomly selected specimens. The specimens were from patients *screened* (i.e. not necessarily enrolled and treated) for the trastuzumab trials. The ratio of CTA positive and negative samples will be in a forced 1:1 ratio. All pathologists and technicians performing FISH will be blinded to results from previous CTA assessments.

Specimens: Archived slides from patients previously screened for trastuzumab studies will be retrieved and accessioned for testing by non-laboratory staff. The number of slides per patient will be tracked. Demographic information will accompany the specimens including patient initials, ID number, date of collection and date of birth.

Blinding: Specimen management staff, the project manager and the project monitor will have password restricted access to the database which contains specimen numbers from previous protocol (CTA) along with newly assigned specimen numbers for this protocol.

Assays:

The CTA is an avidin-biotin method of IHC using both the 4D5 and CB11 antibodies. Results were reported on a 0-3+ scale with "positive" meaning more than 10% of the cells have weak to moderate complete membrane staining stain (2+) or moderate to strong staining (3+).

The FISH assay is PathVysion HER-2/neu DNA Probe manufactured by Vysis or INFORM HER2 manufactured by Ventana [INFORM was ultimately not used in this study.] The probe is a 190 Kb SpectrumOrange directly labeled fluorescent DNA probe specific for the HER-2/neu gene locus at 17q11.2-q12. The CEP 17 DNA probe is a 5.4 Kb Spectrum Green directly labeled specific for the alpha satellite

DNA sequence at the centromeric region of chromosome 17 at 17p11.1-q11.1. The scoring is based on the ratio of the number of HER-2/neu to CEP 17 signals per nucleus. A signal ration of ≥ 2.0 is considered abnormal amplification. Assessment will be of 60 interphase tumor cell nuclei per target.

Data collection: Results of FISH determination will be recorded manually and entered into a secure database. Test result data will be supplied in electronic format to Genentech upon completion of each of the Vysis and Ventana assays or more often as requested.

Data analysis: The proportion of samples sharing agreement between the CTA and FISH will be calculated and the confidence intervals will be obtained. This will be called the level of concordance between the two assays. Data will be presented in a 2x2 table [FISH (+) and FISH (-) vs IHC (+) and IHC (-) where IHC (+) = score of 2+ or 3+]. A one-sided test on proportion will be conducted on the concordance rate at the 5% significance level. The secondary analysis will extrapolate the results to the population using the population distribution of 0, 1+, 2+ and 3+ scores. The point estimate of the concordance level of the population will be calculated accordingly with the target being 85%-90% range. The Kappa statistic for the 2x2 table will be calculated. Exploratory analysis to identify and optimal cut point for FISH ratio denoting positivity will be assessed using 1.5, 2.0 and 2.5 as possible cut points. Additional cut offs may be assessed.

C. Clinical Outcome Study

There was no prospective protocol for the clinical outcome study. There were only standard operating procedures (SOPs) written by LabCorp and Press for running assays on tissue samples. At the time of the writing of this document FDA is awaiting receipt of these SOPs from the sponsor.

Press Protocol: Laboratory Testing for HER-2/neu Gene Amplification: Retrospective Study of Genentech Herceptin Clinical Trial Specimens.

Primary Endpoint: Test approximately 345 paraffin-embedded breast cancer tissue specimens for HER-2/neu gene amplification using FISH and report results to Genentech, Inc. for comparison with clinical outcome data already collected in three clinical trials (H0648g, H0649g, and H0650g).

Study Design: Retrospective, single center study to test tissue samples for HER-2/neu gene amplification.

Specimens: Paraffin-embedded tissue specimens currently in storage at LabCorp, Inc.

Assays: not described

Data collection: not described

Data analysis: not described

D. LabCorp vs Press Validation Study

Not included in application. Awaiting protocol outline from sponsor.

IV. RESULTS OF FISH TESTING ON ARCHIVED TISSUE SAMPLES FROM CLINICAL TRIALS

Data submitted were assessed for the following items:

- Difference in methods between the two FISH testing laboratories (LabCorp and Press)
- Failure rate of FISH (i.e. the relative number of times a FISH result could not be obtained when tissue was available for testing)
- Difference in performance of FISH between the two FISH testing laboratories
- Concordance between FISH results and previously conducted CTA results
- Degree of agreement between LabCorp and Press laboratories
- Retrospective clinical outcome assessment by FISH result and IHC results

A. Methodological Differences Between LabCorp and Press

Due to a variety of tissue specific factors (e.g. storage, age of samples, previous staining), both laboratories conducted the FISH assays with some modifications to the PathVysion package insert instructions. Table 1 identifies the methodologic procedures for slide preparation and staining which differed between each lab and that recommended in the PathVysion package insert.

Table 1. Differences in methodology between two laboratories performing FISH.

Parameter	LabCorp	Press Lab	PathVysion PI
Specimen fixation	Formalin-fixed, paraffin embedded	Formalin-fixed, paraffin embedded	Formalin-fixed, paraffin embedded
Specimen slide prep	4-6 μ m sections not coverslipped	4-6 μ m sections with Cytoseal mounting and coverslip	4-6 μ m sections
Specimen storage	\geq 2.5 years	\geq 2.5 years	Length not specified
Removal of coverslip	Not applicable	Xylene 24-48hr followed by EtOH to remove xylene	Not specified
Chaotrope Treatment (part of deparaffinization)	1M NaSCN 80°C for 60 min	1M NaSCN 80°C for 30 min	1M NaSCN 80°C for 30 min
Processing	VP 2000 (80%) and manual (20%)	VP 2000 (100%)	Not addressed
Probe application	10-20 μ l for 24 x 40 coverslip; amount of probe sometimes increased for larger coverslip	10 μ l probe for 22x22 coverslip; amount of probe doubled for larger coverslip	10 μ l probe for 22x22 coverslip

Counterstain	10 μ l DAPI for 24 x 40 coverslip	10 μ l DAPI for 22x22 coverslip	10 μ l DAPI for 22x22 coverslip
Microscope filters	DAPI/Or/Gr v.2 DAPI/Or v.2 Green single	DAPI/Or/Gr v.2 DAPI/Or v.2 DAPI/Gr	DAPI/9-Or DAPI/Gr DAPI/Or/Gr
Microscope Objectives	10x, 40x, 100x oil	5x, 10x, 20x, oil:40x, 63x, 100x	40x, 63x, 100x
Number of nuclei scored	40 if ratio 1.8-2.2 then 60	60	60
Signal enumeration	Not specified	Per PI	Guidelines provided
Ratio calculation (HER2:CEP17)	Total HER2/Total CEP 17; for high number of signals a mid range chart was used.	Average HER2/Average CEP17	Not specified
Control slides	Formalin fixed, paraffin-embedded breast cancer tissue previously identified as HER2 positive by IHC or FISH. Performance characteristics not specified.	Internal control (benign and normal cells and non-amplified and amplified cell lines. If negative control failed normal lymphocytes used. ProbeChek HER2. At least 2 controls with successful assay results	ProbeChek control slides

B. Success/Failure Rates for Tissue Samples tested by FISH

Summary of Sample Testing

For the CTA vs FISH concordance study, the prospective plan was to test 600 out of 5251 unique patients with available tissue from 5998 patients screened. Selection of samples was random, but in a forced 1:1 ratio. Instead, 623 samples were tested of which 49% were 0-1+ and 52% were 2-3+ by CTA. Of the 623 samples, a FISH result was obtained on 529 samples for a testing failure rate of 15% by LabCorp (Diagram 1).

For the Clinical Outcome study, the plan was to test all available samples (784 out of 799 patients enrolled). Patients with at least 2 unstained slides numbered 618 and of

these, a result was obtained on 540 samples for a testing failure rate of 13% by LabCorp. The total missing data for the clinical trials was 32%. To minimize the amount of missing data, Press tested 244 samples for which there were < 2 unstained slides or no FISH result. Of these, a result was obtained on 225 samples for a testing failure rate of 8% by Press. (Diagram 2)

For the LabCorp vs Press Validation study, the plan was to test 90 samples in a forced 1:1 ratio of FISH (+) vs (-). Due to poor concordance with LabCorp negative specimens, additional samples were tested: 52 known (+) and 108 (-). Of the 250 samples tested by Press, a result was obtained on 223 samples for a testing failure rate of 11% by Press. (Diagram 3)

**Diagram 1: Sample testing flow diagram for CTA vs FISH
Concordance Study**

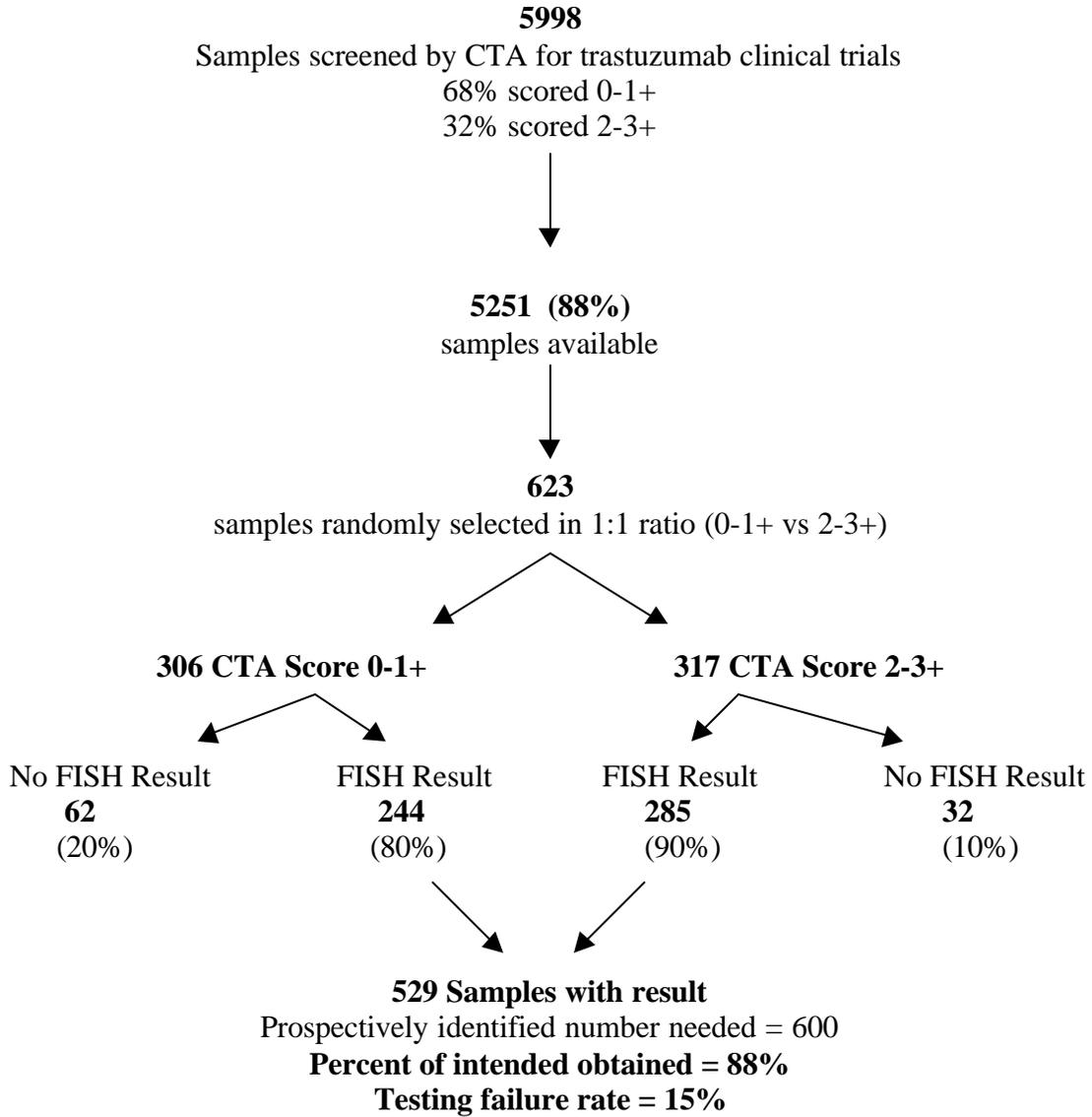


Diagram 2: Sample testing flow diagram for Clinical Outcome Study

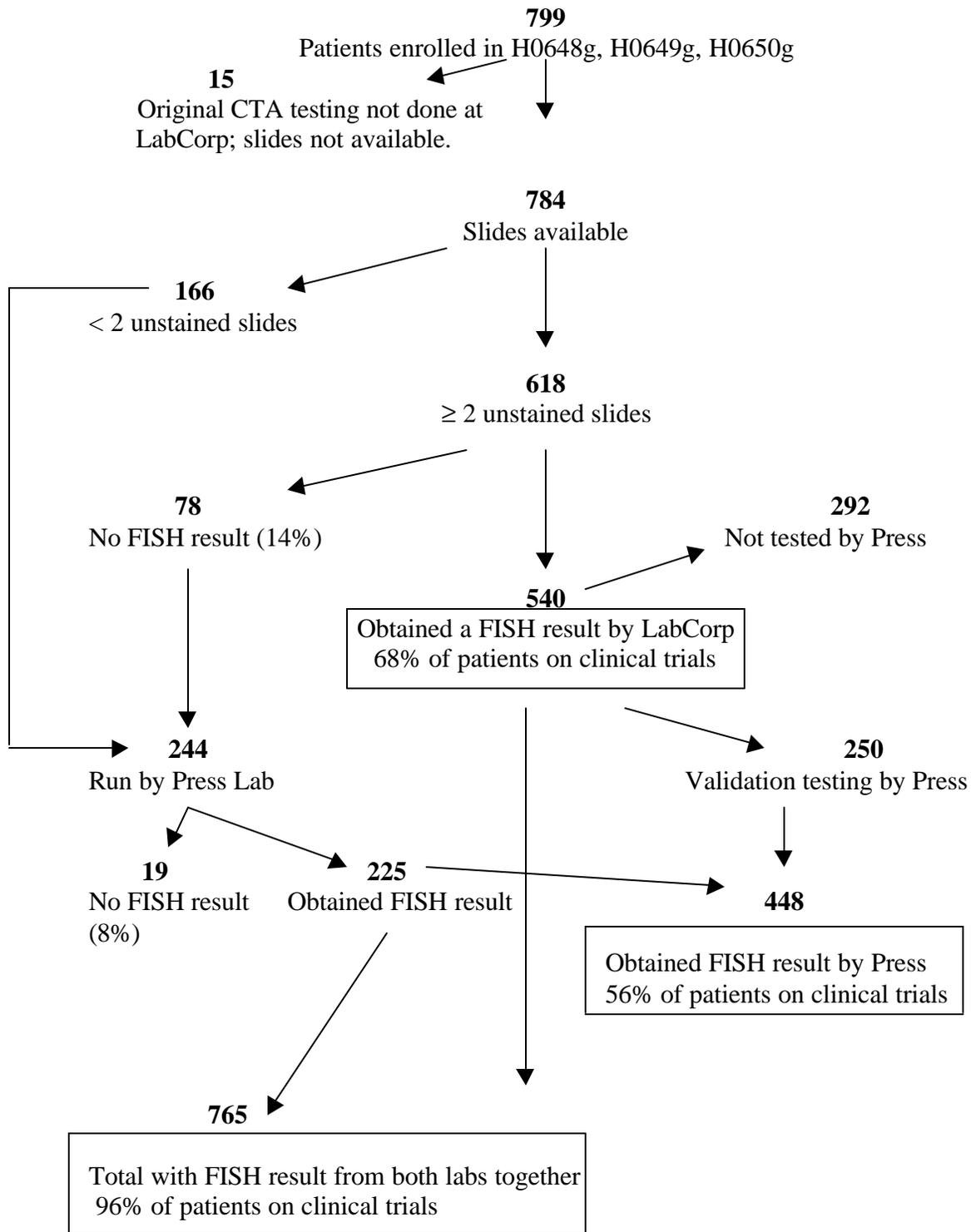


Diagram 3: Sample testing flow diagram for LabCorp vs Press Validation Study

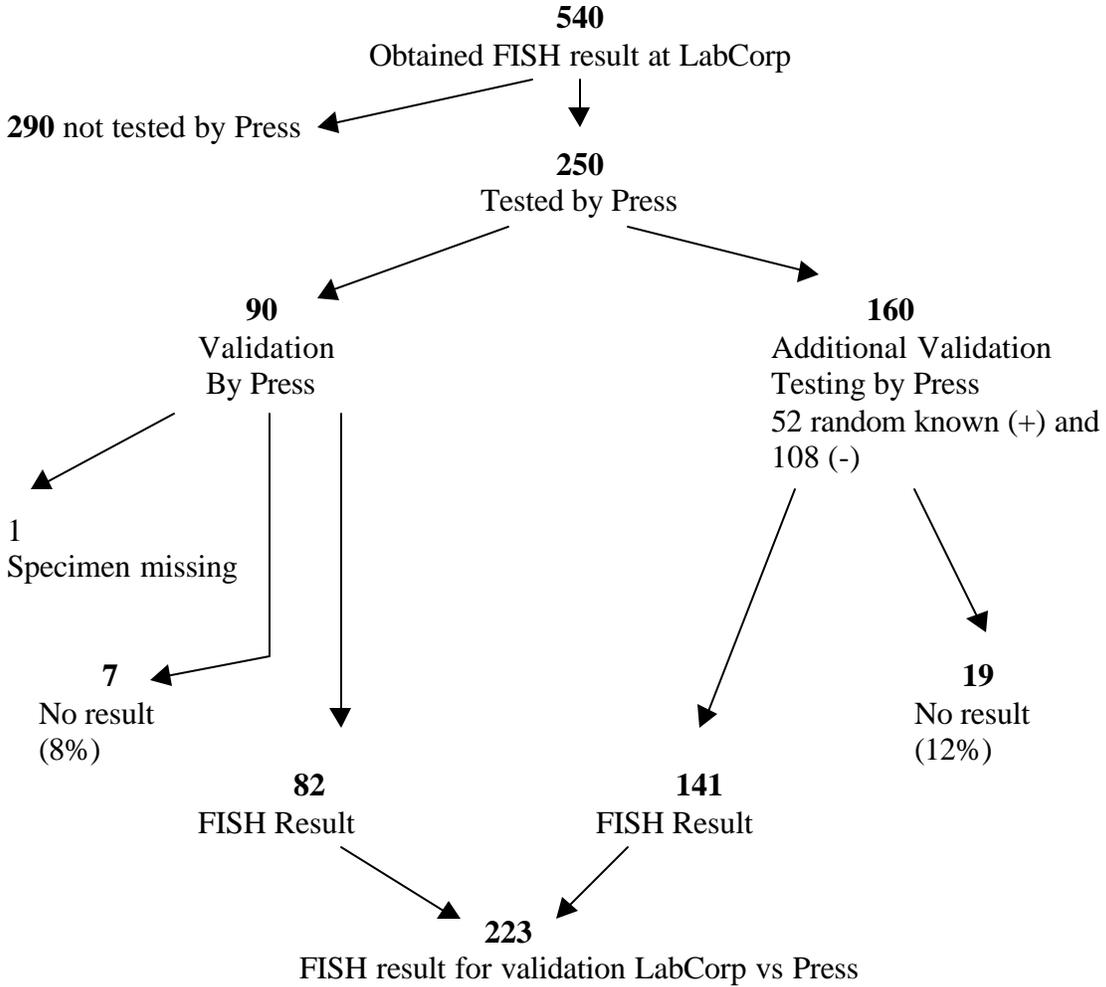
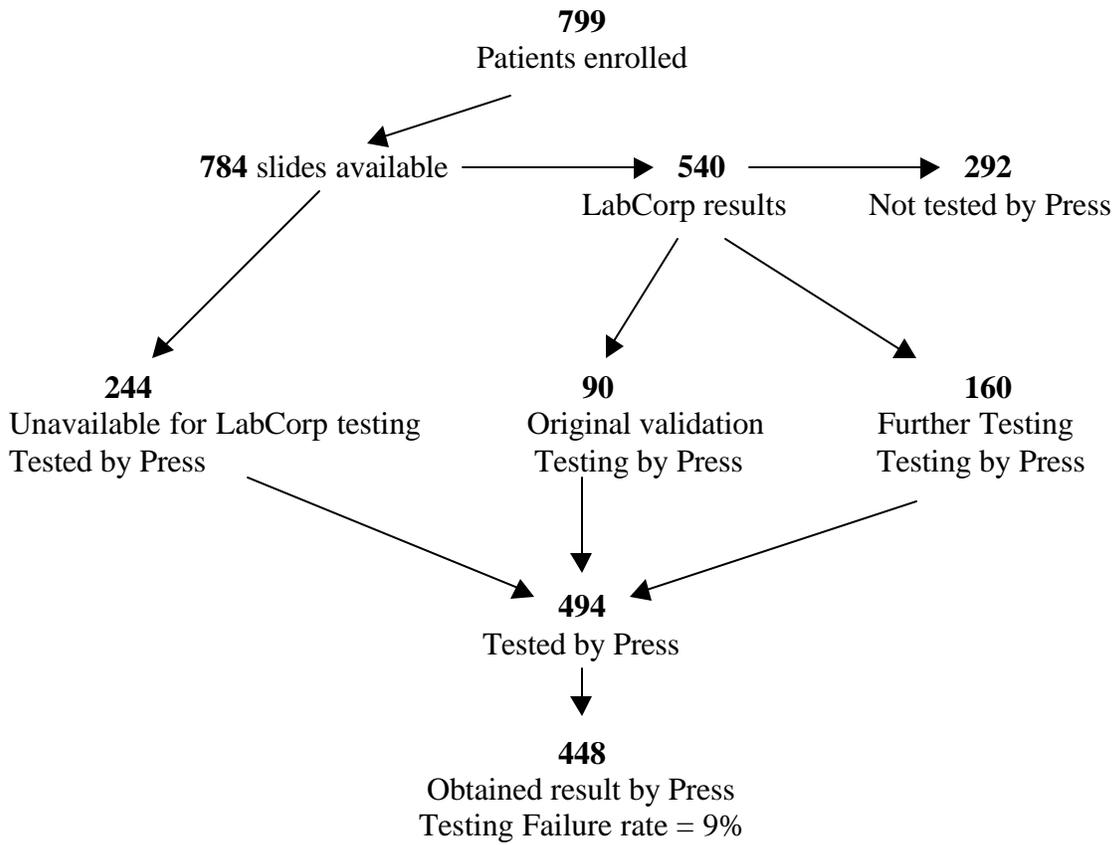


Diagram 4: Sample testing flow diagram for Press lab



C. Comparison of FISH Results Between LabCorp and Press Laboratories

FISH scoring required the counting of nuclei for CEP 17 staining (centromere of Chromosome 17) and HER2 staining. The ratio of the HER2:CEP 17 was the final score. Scores ≥ 2.0 were considered positive for HER2 amplification.

Case Report Form Review

All case report forms (CRF's) for samples from study H0648g which were tested at both laboratory sites were reviewed. Case report forms were reviewed for the number of readers per laboratory site, accuracy of calculations, and for comparison of centromere CEP 17 and HER2 scores between LabCorp and Press. It was noted that at LabCorp, there were scores sheets completed by 3 different individuals, however, 90% of these were completed by one of the three. At the Press lab, all scores sheets were completed by Dr. Press together with one of two other individuals in his laboratory. Approximately 5% of CRF's were reviewed for accuracy of calculations and all were accurate. Each lab used a different method for calculating the final FISH score.

Mean centromere counts and HER2 counts were calculated by the reviewer from LabCorp CRF's; mean centromere counts and HER2 counts were entered on Press CRF's. Calculated ratios of HER:centromere were entered on both sets of CRF's. Comparisons were made between the scoring of patient samples. The data were not provided as SAS tables and therefore, paired correlation analyses could not be performed within the time frame of this review.

There was a consistent trend for the scores at LabCorp to be lower overall than those done by the Press lab. Of the 198 patient cases reviewed, a result was obtained by both labs for 141 cases. The remaining 57 cases had a failed result at one or both labs. Mean numeric scores were lower overall at LabCorp than at the Press lab for centromere and particularly for HER2 scores: centromere 2.4 vs 2.5 and HER2 5.2 vs 11.9, LabCorp vs Press respectively. As a result the ratio scores were also lower at LabCorp 83% (117 out of 141) of the time with mean values of 2.4 vs 4.7, LabCorp vs Press respectively.

There were 23 cases in which both labs obtained discordant results (16%). Of these, in only one case was the LabCorp result positive while the Press result was negative; the other 22 cases were negative by LabCorp and positive by Press. The mean value comparisons for these 23 cases are as follows: centromere 3.5 vs 3.2, HER2 5.8 vs 15.2, and ratio 1.7 vs 5.1, LabCorp vs Press respectively.

Exploratory analyses of the concordance data appear in Appendix A. These include a concordance analysis of LabCorp vs. Press, exploration of alternative cut off points for FISH score, and concordance of numeric FISH with numeric IHC.

D. CTA vs FISH Concordance Study

FISH results were compared with those obtained using the Clinical Trial Assay IHC testing. The pre-specified comparison was FISH (+) and FISH (-) vs IHC (+) and IHC (-) where IHC positive included scores of 2+ and 3+ (Table 2). The current package insert for trastuzumab, however, states that “Data from both efficacy trials suggest that the beneficial treatment effects were largely limited to patients with the highest level of HER2 protein overexpression (3+).” For that reason FDA also undertook an exploratory analysis where IHC (+) was defined as 3+ only (Table 2). It is important to consider the outcomes for patients scored as 2+ vs 3+ juxtaposed to those scored as FISH (+) vs FISH (-). The interpretation of IHC 2+ has been a problematic area because the distinction between 1+ vs 2+ and in some cases between 2+ vs 3+ is very difficult and subject to technical differences between laboratories and to reader interpretation. It has also been a matter of debate whether a score of 2+ represents protein overexpression.

Please note that for purposes of the primary analysis, the value obtained at LabCorp took precedence over that obtained at Press in cases where a result at both labs was determined. This was consistent with the prospective plan, since the Press results were collected for the purpose of filling in missing data and not initially for validation of LabCorp results.

The concordance was moderate for the primary analysis (Kappa = 0.64) and stronger when 2+ by IHC was defined as negative (Kappa = 0.80). FISH testing missed 11% of the 3+ samples and selected 4% of the 0-1+ samples as positive. Twenty-four percent of the 2+ samples were FISH (+).

Table 2. Concordance of FISH (LabCorp and Press pooled with LabCorp result taking precedence over Press) vs CTA for screened samples.*

	CTA = 0	CTA = 1+	CTA = 2+	CTA +3+
FISH (+)	7	2	21	176
FISH (-)	207	28	67	21
Unknown	57	5	8	24

*Analysis performed with assumption that CTA 2+ and 3+ are a positive result and CTA 0 and 1+ are a negative result: **Kappa = 0.64 with 95% CI (0.58, 0.70) and Concordance 82% with 95% CI (78%, 85%)**. This is in agreement with the sponsor assessment.

*Analysis performed with assumption that CTA 3+ is positive result and CTA 0, 1+, and 2+ are negative results: **Kappa = 0.80 with 95% CI (0.74, 0.85) and Concordance = 90% with 95% CI (88%, 93%)**. Not performed by sponsor.

Another exploratory analysis undertaken by FDA was to look at concordance of the CTA with FISH for all samples assessed with a FISH result from patients enrolled on the three clinical trials conducted. These data do not assess concordance with 0 and 1+ samples. For this analysis an IHC positive score was defined as 3+ only (Table

3). Of those patients tested by CTA as 3+, FISH testing missed 13% (75 out of 559 which were 3+). Of those which were 2+, FISH was positive in 34% (65 out of 190 which were 2+). The level of concordance (kappa = 0.52) is not as good when looking at this selected population of patients.

Table 3. Concordance in clinical trial samples of enrolled patients from all studies (h0648, 649 and 650g) with pooled FISH data.*

	CTA = 0	CTA = 1+	CTA = 2+	CTA +3+
FISH (+)			65	484
FISH (-)		1	125	75

*Analysis performed with assumption that CTA 3+ is positive and CTA 0, 1+, and 2+ are negative result: Kappa = 0.52 with 95% CI (0.45, 0.59) and Concordance = 81% with 95% CI (79%, 84%).

The HercepTest package insert provides additional information which should be considered along with these results. Concordance testing comparing HercepTest to CTA was conducted with a separate bank of tumor specimens (i.e. not from the trastuzumab clinical trials). In another study, gene amplification was compared to HercepTest results on tumor specimens previously characterized using Northern, Southern, and Western blotting, FISH, and IHC. Please refer to HercepTest package insert in Appendix C.

E. Clinical Outcome Study

The clinical outcome data discussion will be limited to samples tested from studies H0648g and H0649g. The H0650g study was not accepted by FDA for efficacy analysis because the design of the trial was felt to be an inappropriate representation of patients with primary metastatic breast cancer since it enrolled patients who either refused chemotherapy or could not receive chemotherapy. Because the efficacy data have not been rigorously reviewed by FDA the clinical outcome data for purposes of FISH analysis are not represented here.

It is important to note that there are no clinical outcome data for patients who were IHC 0-1+ and FISH (+) or FISH (-) because by definition all patients were 2+ or 3+ in order to be eligible for these trials.

For purposes of discussion, a variety of subgroups will be discussed in the clinical outcome results. Terminology will be defined as follows:

- 3+ = Patients whose tumors scored as 3+ on IHC testing using CTA
- 2+ = Patients whose tumors scored as 2+ on IHC testing using CTA
- FISH (+) = Patients whose tumors scored as FISH \geq 2.0 using PathVysion
- FISH (-) = Patients whose tumors scored as FISH $<$ 2.0 using PathVysion
- FISH+/3+ = Patients whose tumors scored as 3+ and FISH \geq 2.0

- FISH+/2+ = Patients whose tumors scored as 2+ and FISH ≥ 2.0
- FISH-/3+ = Patients whose tumors scored as 3+ and FISH < 2.0
- FISH-/2+ = Patients whose tumors scored as 2+ and FISH < 2.0

A prospective analysis plan did not exist for this study. All analyses presented are considered exploratory. The clinical outcome study analysis was also limited by the fact that for the H0648g study treatment subgroups (anthracycline with or without trastuzumab, paclitaxel with or without trastuzumab) the number of patients was very small. For that reason, only the pooled groups (chemotherapy with or without trastuzumab) are presented. Even with that, some of the subgroups such as FISH+/2+ are very small making interpretation of results difficult and very limited.

Finally, data analyses of 3+ vs 2+ are also included here for comparison. As noted in the trastuzumab label, "Data from both efficacy trials suggest that the beneficial treatment effects were largely limited to patients with the highest level of HER2 protein overexpression (3+)."

Presented here are clinical outcome data for H0648g and H0649g. The primary endpoints for the two clinical trials were median time to progression on the H0648g study and overall response rate for the H0649g study. The secondary endpoints for the H0648g study presented here include overall response rate, and median overall survival. Because the H0649g study was a single arm study, time to progression and overall survival will not be presented here as there is no comparative arm.

Study H0648g

Time to Progression

Time to progression Kaplan-Meier analysis of H0648g appears below for patients whose tumors tested as FISH(+), FISH(-), 3+, and 2+ (Figures 1-4). For additional exploratory analyses of subgroups (e.g. FISH+/3+, FISH-/3+, etc) please see Appendix A.

Figure 1. **Time to Progression, FISH (+) samples**, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.44 (0.34, 0.57) and the p value < 0.001. This is in agreement with the sponsor assessment.

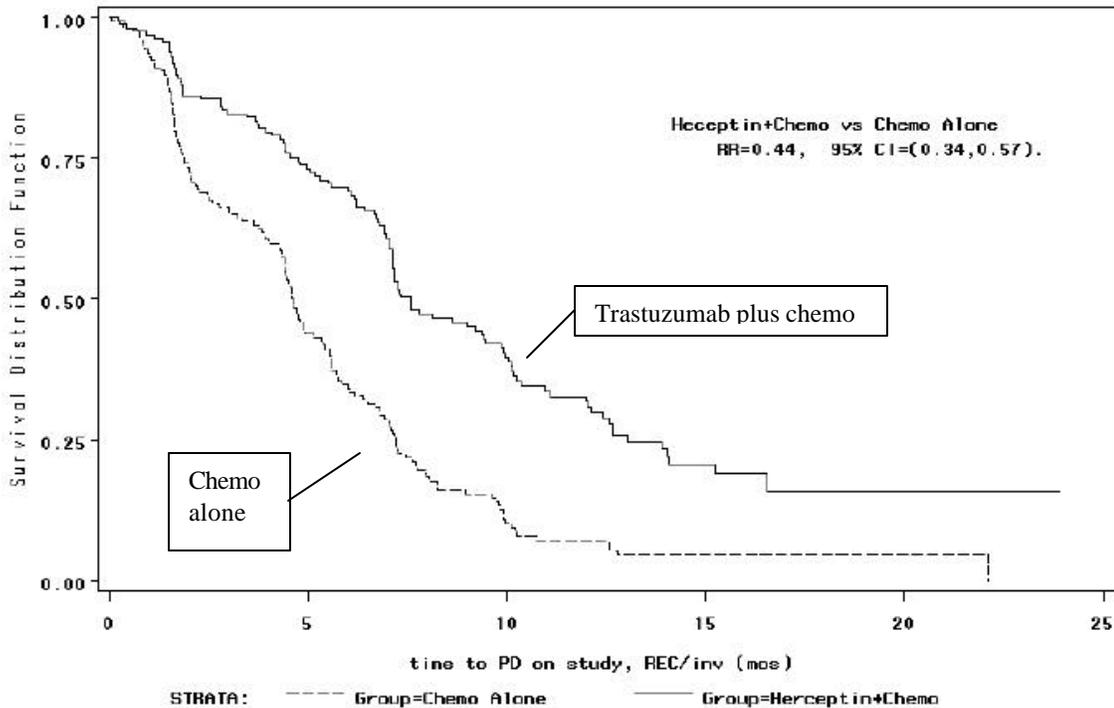


Figure 2. **Time to Progression, FISH (-) samples**, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.66 (0.45, 0.99) and the p value is 0.043. This is in agreement with the sponsor assessment.

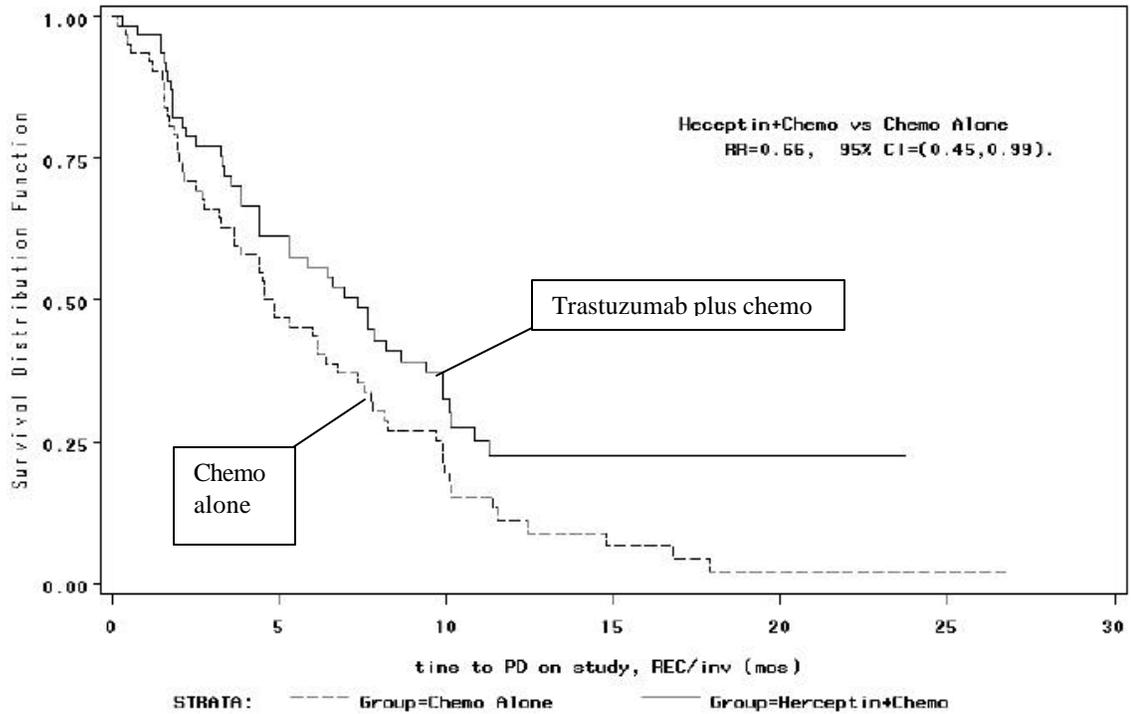


Figure 3. **Time to Progression, 3+** samples (regardless of FISH status) for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.42 (0.33, 0.55) and the p value is < 0.001.

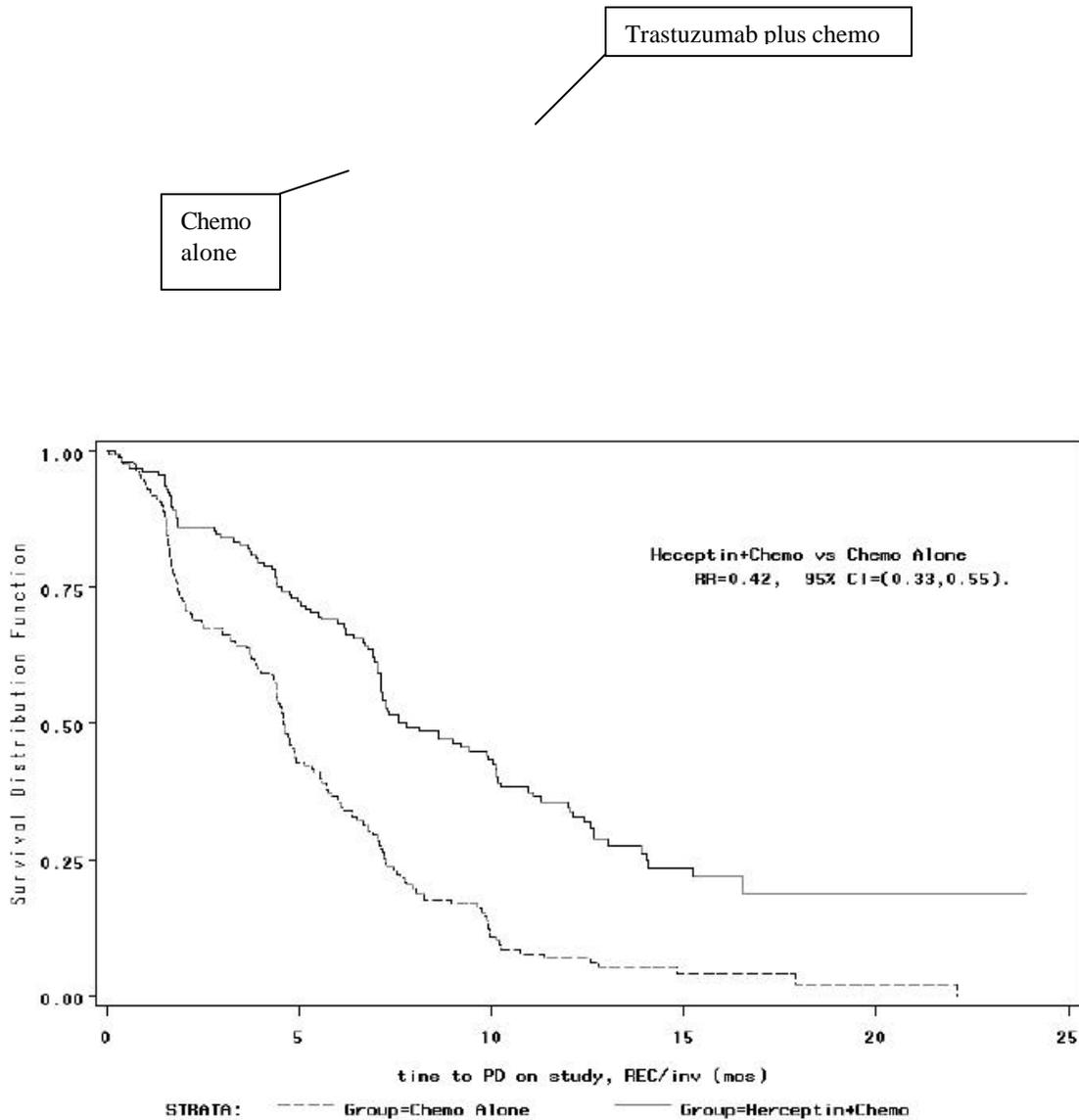
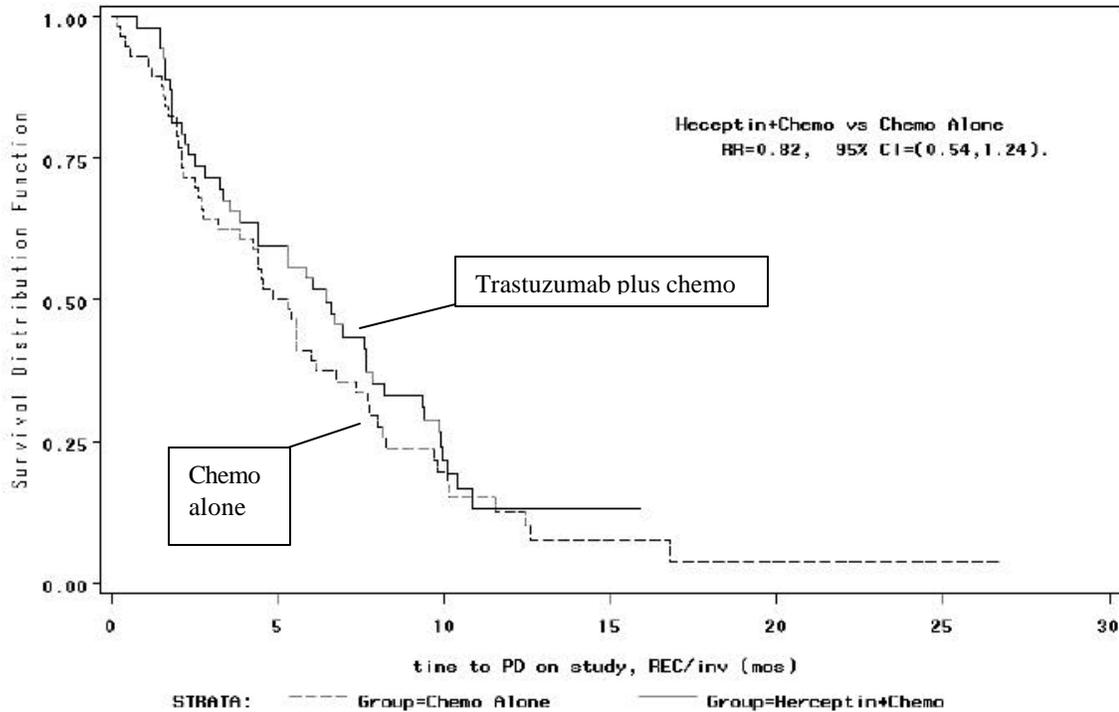


Figure 4. **Time to Progression, 2+** samples (regardless of FISH status) for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.82 (0.54, 1.24) and the p value is 0.34.



Reviewer comment: For time to progression, FISH (+) patients (regardless of IHC score) and 3+ patients (regardless of FISH score) both had greater clinical benefit from the addition of trastuzumab to chemotherapy than did FISH (-) or IHC 2+ patients respectively. However, FISH (-) patients did appear to experience clinical benefit, though not to as large a degree. This is likely due to inclusion of some FISH-/3+ patients (see exploratory analyses in Appendix A).

Overall Response Rate

Overall response rate was defined as objective complete plus partial responses sustained for at least 4 weeks. Below are the analyses for response rate by FISH result, IHC result and various subgroups. These numbers differ slightly from those in the package insert (e.g. for 2+/3+;T+C, the ORR for patients with FISH result is 114 out of 226 or 50% which is higher than the 45% listed in the PI). The values in the package insert were based on the FDA assessment of responder status and were slightly lower than the sponsor's original report of response. We have not included the FDA assessed values for response here for the following reasons: 1) the data are not markedly different from what appears below, 2) these are all retrospective subgroup analyses from which firm conclusions will be difficult to justify.

Table 4. H0648g: Overall Response Rate by FISH results (+) vs (-) (Response is defined as complete plus partial response sustained for ≥ 4 weeks.) Sponsor analysis.

	2+/3+ T+C n = 226	2+/3+ C n = 225	3+ T+C n = 169	3+ C n = 167	2+ T+C n = 57	2+ C n = 58
FISH (+) n = 325	54% (89/164)	30% (49/161)	55% (81/148)	28% (40/145)	50% (8/16)	56% (9/16)
FISH (-) n = 126	40% (25/62)	38% (24/64)	62% (13/21)	55% (12/22)	29% (12/41)	29% (12/42)

T+C = Trastuzumab plus chemo
C = Chemo alone

Table 5. H0648g: Overall Response Rate by IHC score 2+ vs 3+ from the trastuzumab package insert. (Response is defined as complete plus partial response sustained for ≥ 4 weeks.)

	2+/3+ T+C n = 235	2+/3+ C n = 234	3+ T+C n = 176	3+ C n = 173	2+ T+C n = 59	2+ C n = 61
Any FISH N = 469	45% (106/235)	29% (67/234)	49% (87/176)	27% (46/173)	32% (19/59)	34% (21/61)

T+C = Trastuzumab plus chemo
C = Chemo alone

Reviewer comment: For overall response rate, FISH (+) patients (regardless of IHC score) and 3+ patients (regardless of FISH score) both had greater clinical benefit from the addition of trastuzumab to chemotherapy than did FISH (-) or IHC 2+ patients respectively. Patients in the FISH-/3+ group may have also benefited, though the number of patients is too small to draw firm conclusions.

Overall Survival

Overall Survival Kaplan-Meier analysis of H0648g appears below for patients whose tumors tested as FISH(+), FISH(-), 3+, and 2+ (Figures 5-8). For additional exploratory analyses of subgroups (e.g. FISH+/3+, FISH-/3+, etc) please see Appendix A .

Figure 5. **Overall Survival, FISH (+) samples**, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.69 (0.53, 0.91) and the p value is 0.007. This is in agreement with the sponsor assessment.

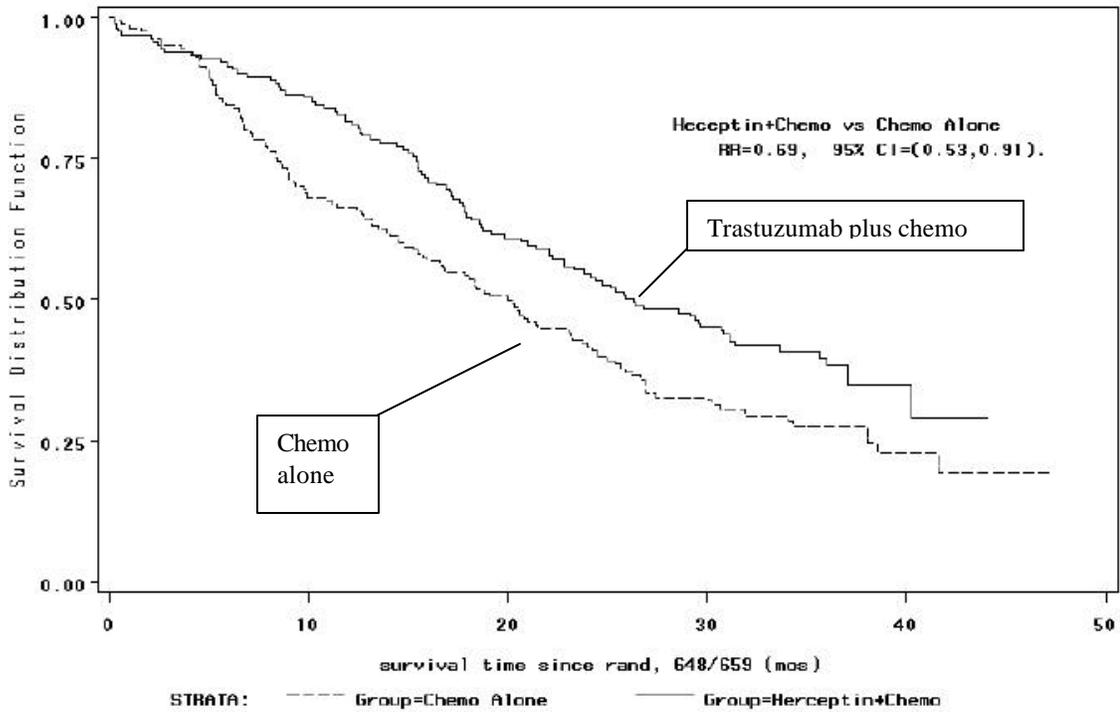


Figure 6. **Overall Survival, FISH (-)** samples, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 1.07 (0.70, 1.63) and the p value is 0.76. This is in agreement with the sponsor assessment.

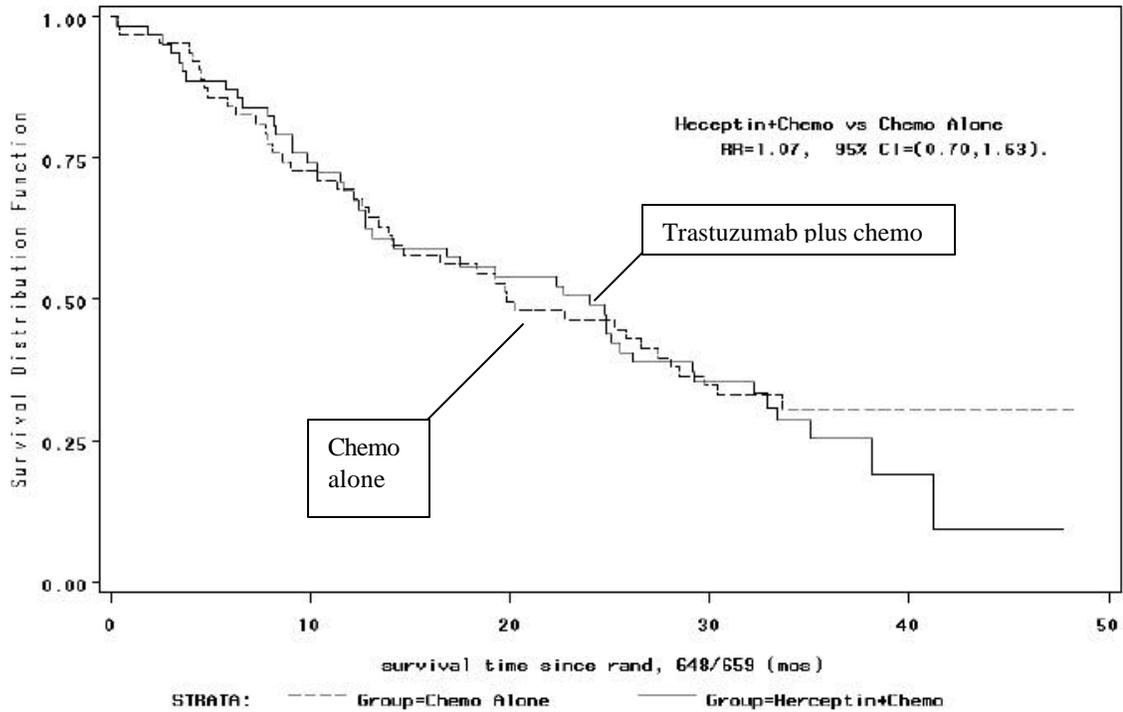


Figure 7. **Overall Survival, 3+** samples (regardless of FISH status) for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.70 (0.54, 0.92) and the p value is < 0.010.

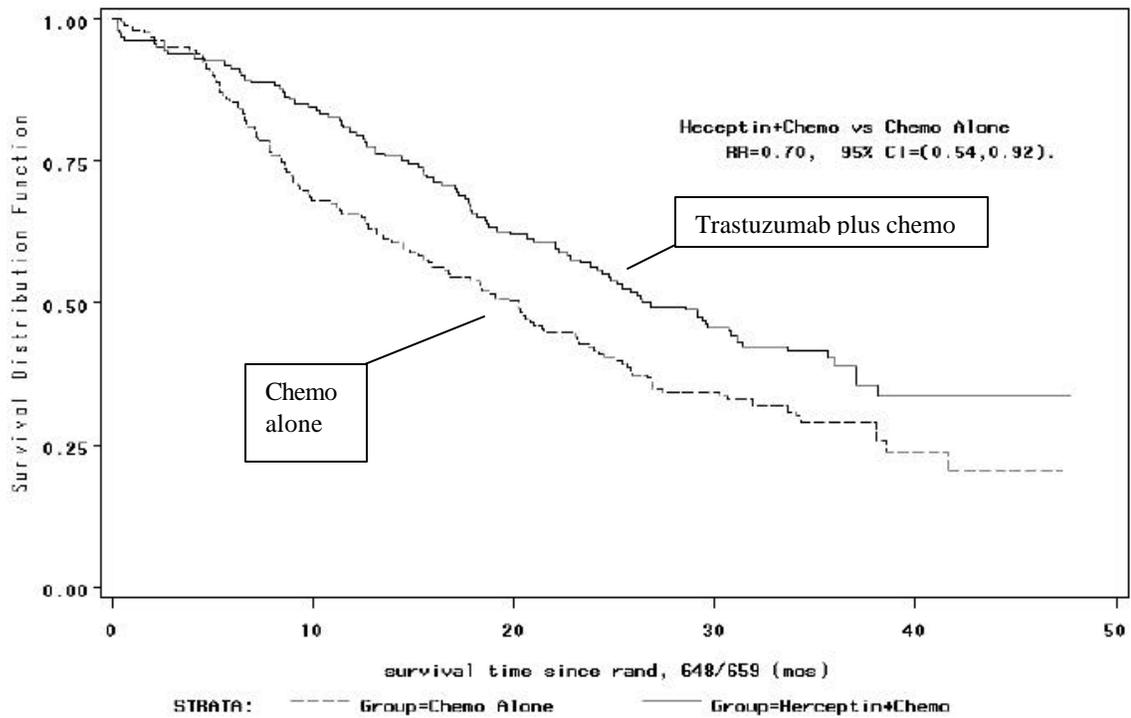
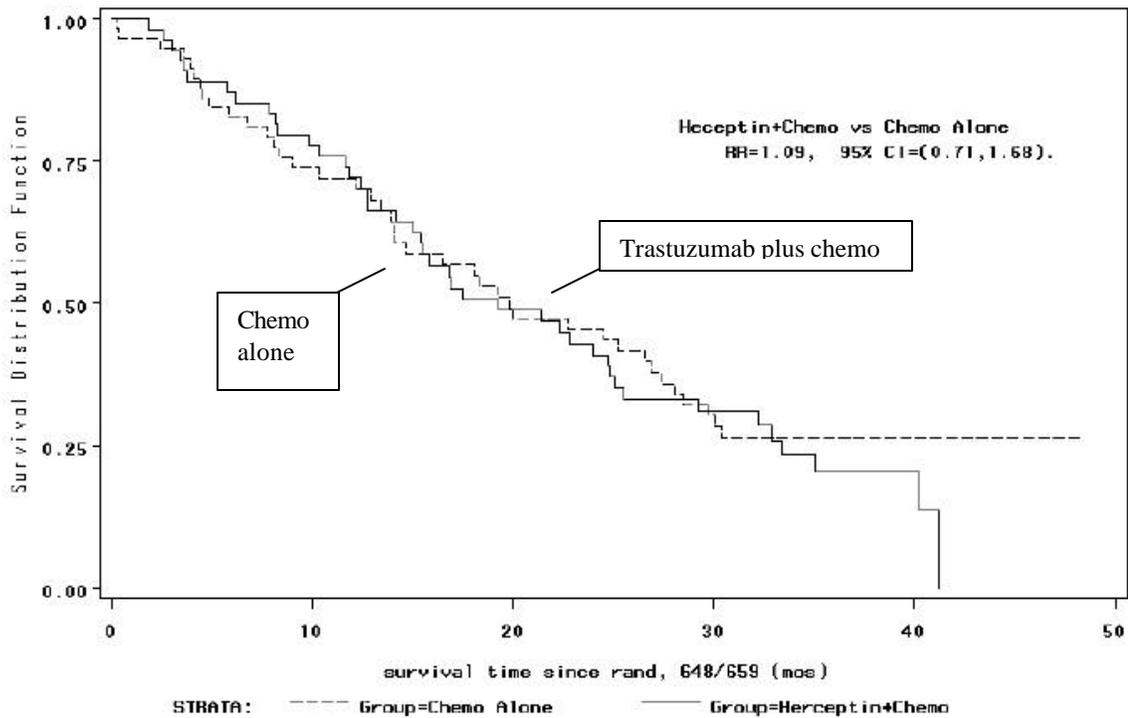


Figure 8. **Overall Survival, 2+ samples** (regardless of FISH status) for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 1.09 (0.71, 1.58) and the p value is 0.70.



Reviewer comment: For overall survival, FISH (+) patients (regardless of IHC score) and 3+ patients (regardless of FISH score) both had greater clinical benefit from the addition of trastuzumab to chemotherapy than did FISH (-) or IHC 2+ patients respectively.

Study H0649g

Overall Response Rate

Table 6. H0649g: Overall Response Rate by FISH results (+) vs (-) (Response is defined as complete plus partial response sustained for ≥ 4 weeks.) Sponsor analysis.

	CTA 3+	CTA 2+	ALL
FISH (+)	22% (30/136)	11% (3/27)	20% (33/163)
FISH (-)	0% (0/21)	0% (0/25)	0% (0/46)
ALL	19% (30/157)	6% (3/52)	

Reviewer comment: For overall response rate, FISH (+) patients (regardless of IHC score) and 3+ patients (regardless of FISH score) both had greater clinical benefit from the addition of trastuzumab to chemotherapy than did FISH (-) or IHC 2+ patients respectively. Given the small number of patients in the CTA 2+ subgroup, it is difficult to conclude differences between FISH-/2+ and FISH+/2+ patients.

V. CONCLUSIONS

- A. FISH and IHC testing are both semiquantitative assays in the most optimal circumstances. Even in this setting, there can be considerable variability in laboratory methodology as evidenced by systemic differences between the two laboratories conducting the FISH assays for these studies. LabCorp FISH scores were consistently lower than those of the Press lab, leading to less confidence in scores surrounding the cutoff of 2.0.
- B. There is moderate concordance (82% [78%, 85%] with Kappa statistic = 0.64) between the original CTA assay conducted at LabCorp and the FISH assay conducted at LabCorp using the PathVysion HER2 gene amplification FISH kit. In exploratory analyses, when the CTA positive definition was changed to 3+ only, instead of 3+ and 2+, the agreement between the CTA and FISH remained moderate to good. This was upheld when looking at the agreement between FISH and the original CTA value of enrolled patients.
- C. The level of concordance between CTA and FISH in this study is similar to that observed between HercepTest and CTA in a previous study.
- D. The clinical outcome data does not provide information on the clinical outcome for patients whose tumors score IHC 0 or 1+ and FISH (+). For this reason, there are insufficient data to describe the predictability for response to trastuzumab based upon FISH test results.
- E. The clinical outcome data were not obtained in a fashion to allow direct comparison with IHC, therefore, comparative claims of equivalence cannot be made, and claims of superiority cannot be made. It may be possible to note that trends proceed in a similar fashion for IHC 3+ and FISH(+) subgroups.
- F. The trend toward beneficial effects of trastuzumab in the clinical outcome study are prominent in the FISH (+) subgroup. A smaller clinical benefit was also seen in the FISH (-) subgroup as evidenced in the analysis of time to progression and overall response rate; this did not translate into increased survival, however.
- G. The general magnitude of the beneficial effects of trastuzumab therapy in the FISH (+) subgroup is similar to that in the ICH 3+ subgroup.
- H. There is insufficient information to draw conclusions regarding potential sequencing of HER2 detection assays. For example, there is insufficient data to claim that if IHC is performed first and the score is $\leq 2+$, then FISH should be performed next.

- I. There is insufficient information to draw conclusions regarding the use of FISH and IHC simultaneously. For example, should patients whose tumors are FISH+/2+ receive trastuzumab therapy? Or FISH+/0-1+?
- J. Given the subjective nature of both FISH and IHC testing and given the relative lack of knowledge regarding the functioning and detection of the trastuzumab target, the obtaining of clinical outcome data remains the only gold standard for determination of the ability of a HER2 detection methodology to optimally select patients for trastuzumab therapy.
- K. These retrospective examinations of archived clinical trial samples are useful for hypothesis generation for future studies: Will FISH-/3+ or FISH+/2+ patients benefit from trastuzumab therapy? Is 3+ or FISH + a better predictor of clinical outcome with trastuzumab therapy? Will FISH+/0-1+ patients benefit from trastuzumab therapy?

VI. Appendix A

Exploratory Analyses of Concordance Data

Concordance between LabCorp and Press results

The sponsor conducted a validation study to compare the results between the two laboratories. After the first 90 samples were assessed, there appeared to be underscoring at LabCorp which led to the decision to test additional samples. The effect persisted. Concordance between LabCorp and Press results were assessed by FDA and appear in Table 7. There was poor concordance between LabCorp and Press FISH (+) results. Of the samples testing positive at the Press lab, 32% (37/116) were negative at LabCorp.

Table 7. Concordance between LabCorp and Press Lab Results (H0648g, H0649g, H0650g samples)

	Missing data	Press FISH (+)	Press FISH (-)	Total
Missing data	0	176	50	226
LabCorp FISH (+)	308	79	2	81
LabCorp FISH (-)	14	37	105	142
Total		116	107	223

Cut off point for FISH positivity

Given the consistently lower scoring by LabCorp, it was hypothesized by the reviewer that samples scored as negative by LabCorp with a result in the range of 1.0 to 2.0 would have a high rate of positivity by Press laboratory scoring. The reviewer examined all results from all studies for which there was both a LabCorp result and a Press result. For scores in the range of 1.0 to 2.0 by LabCorp, 38% tested positive by Press. For scores in the range of 1.0 to 2.0 by Press, 1% tested positive by LabCorp.

The reviewer also examined the H0648g study data and the CTA vs FISH concordance data separately. For the H0648g study, 35% of negative LabCorp samples in the 1.0-2.0 range were positive by Press (most of these were 3+ by CTA). For the concordance trial (2+ and 3+ samples only), 29% of 1.0-2.0 LabCorp samples were 3+ by CTA. The effect seemed to be consistent across studies. For the concordance study 0-3+ samples, 10% of 1.0-2.0 LabCorp samples were 3+ by CTA. We estimate that for CTA 2+ and 3+ patients, LabCorp values in the range of 1.0 to 2.0 will miss at least 30% of patients who would benefit from trastuzumab therapy. Likewise for all CTA vales, LabCorp values in the range of 1.0-2.0 will miss 10% of patients who would benefit from trastuzumab therapy based on an IHC score of 3+.

A separate exploratory analysis to explore alternate cut off points for FISH positivity demonstrated 1.0 to be a better cut point than 2.0 in terms of time to progression

clinical outcome data for all three clinical trials. Cut points tested were 1.0, 1.5, 2.0, 2.5, 3.0, and 4.0. The relative risk was best for 1.0 (RR = 0.50) as compared to 2.0 (RR = 0.69), 3.0 (RR = 0.74), and 4.0 (RR = 0.84). This is most likely explained by the consistent lower scoring of the LabCorp slides. While there are many confounding factors in these studies, it is concerning that scores surrounding the cut off and in the range of 1.0 – 2.0 are problematic and may not be as reliable. It is not clear that increasing the number of nuclei counted will overcome this problem. It also does not make biologic sense to have a cut off < 2.0. Therefore, extrapolation of this analysis is problematic.

Numeric FISH score vs CTA result

While it is recognized that the numeric value for the FISH score cannot be interpreted as a quantitative value, we felt that an analysis of scores vs CTA results would be of interest. The data from the concordance study of CTA vs FISH using tissue samples from the screened patient pool run by LabCorp were used in the analysis. The results appear in Table 8. From this distribution, there is a trend for magnitude of CTA score to correlate with magnitude of FISH numeric score. The greatest ambiguity exists for CTA 2+ cases; only 24% of these tested FISH (+) with varying degrees of amplification. High levels of amplification (FISH score > 4) were rare in CTA 0-1+ cases, though the rate of FISH positives in the IHC 0-1+ population (4%) cannot be ignored. Cases which were CTA 3+ were more likely to have high levels of amplification. Any conclusions beyond trend analysis should not be drawn from these data as neither IHC nor FISH assay is meant to be interpreted quantitatively.

Table 8. Concordance analysis of numeric FISH score from LabCorp vs CTA results for samples from screened patient pool.

FISH	CTA Result (num)														
	0			1			2			3			All		
	Row		Col	Row		Col	Row		Col	Row		Col	Row		Col
	N	%	%	N	%	%	N	%	%	N	%	%	N	%	%
Unk	57	61	21	5	5	14	8	9	8	24	26	11	94	100	15
0< - 1	136	72	50	13	7	37	32	17	33	7	4	3	188	100	30
1< - 2	71	53	26	15	11	43	35	26	36	14	10	6	135	100	22
2< - 3	3	12	1	1	4	3	5	19	5	17	65	8	26	100	4
3< - 4	2	6	1	1	3	3	7	21	7	23	70	10	33	100	5
4< - 5	1	3	0				1	3	1	38	95	17	40	100	6
5< - 6							4	15	4	22	85	10	26	100	4
6< - 7							2	6	2	32	94	14	34	100	5
7< - 8	1	4	0				1	4	1	25	93	11	27	100	4
8< - 9							1	14	1	6	86	3	7	100	1
9< - 10										10	100	5	10	100	2
>10										3	100	1	3	100	0
All	271	43	100	35	6	100	96	15	100	221	35	100	623	100	100

Exploratory analyses of Clinical Outcome Data

Study H0648g

Time Progression

Table 9. Median Time to Progression (months) for H0648g subgroups

	3+ ; H+C	3+; C	2+; H+C	2+ ; C
FISH +	7.3	4.6	9.4	5.6
FISH -	11.3	6.3	5.3	4.5

Figure 9. **Time to Progression, FISH+/3+** samples, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.42 (0.32, 0.55).

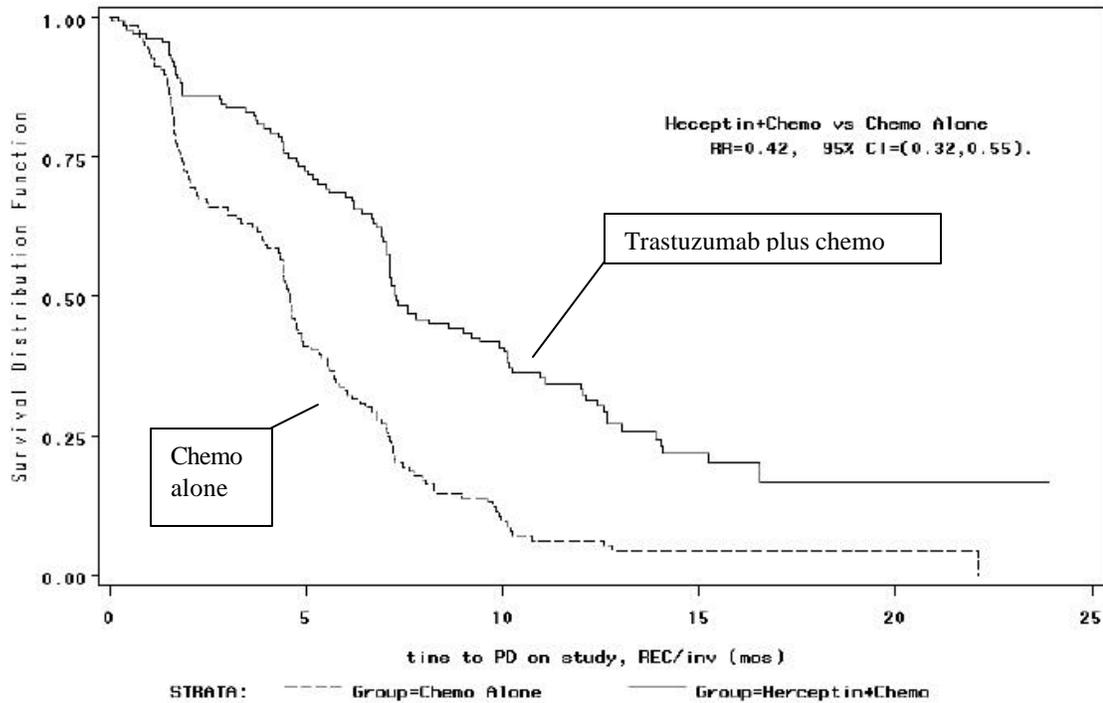


Figure 10. **Time to Progression, FISH-/3+ samples**, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.40 (0.19, 0.87).

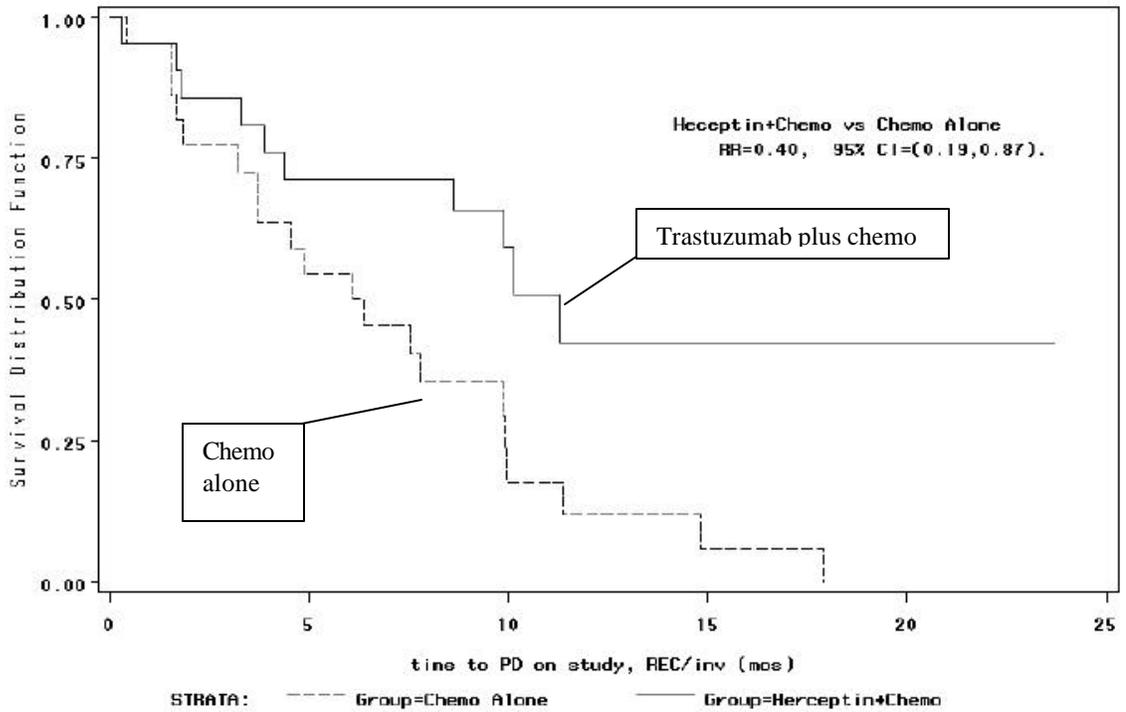


Figure 11. **Time to Progression, FISH+/2+** samples, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.72 (0.31, 1.64).

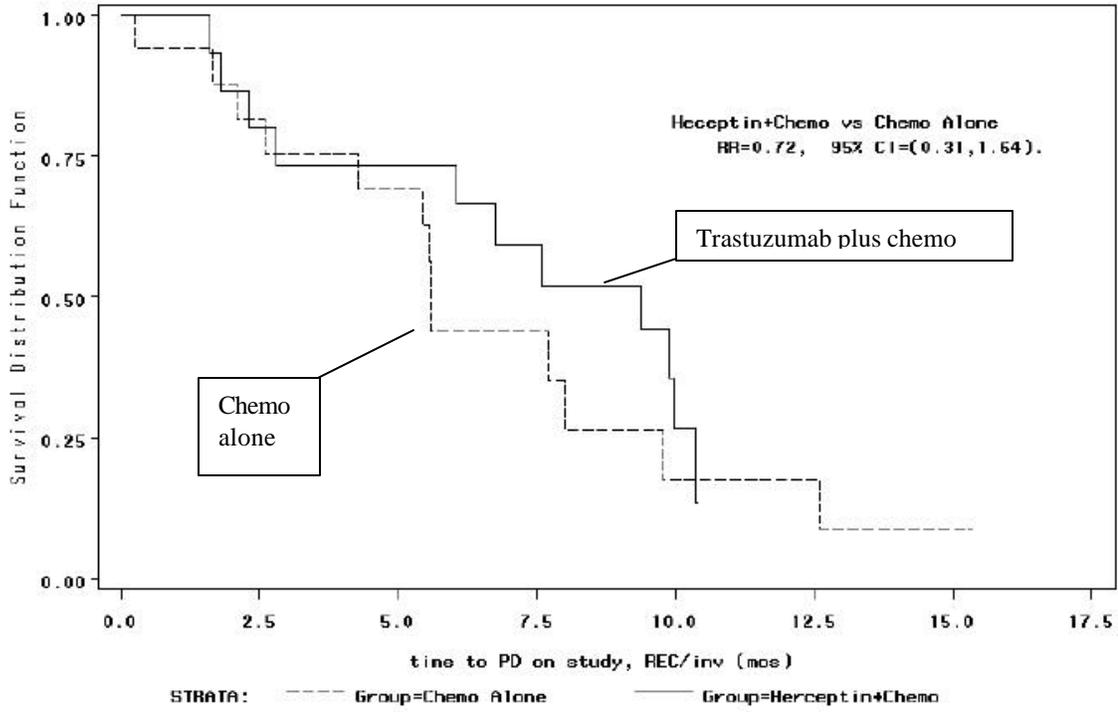
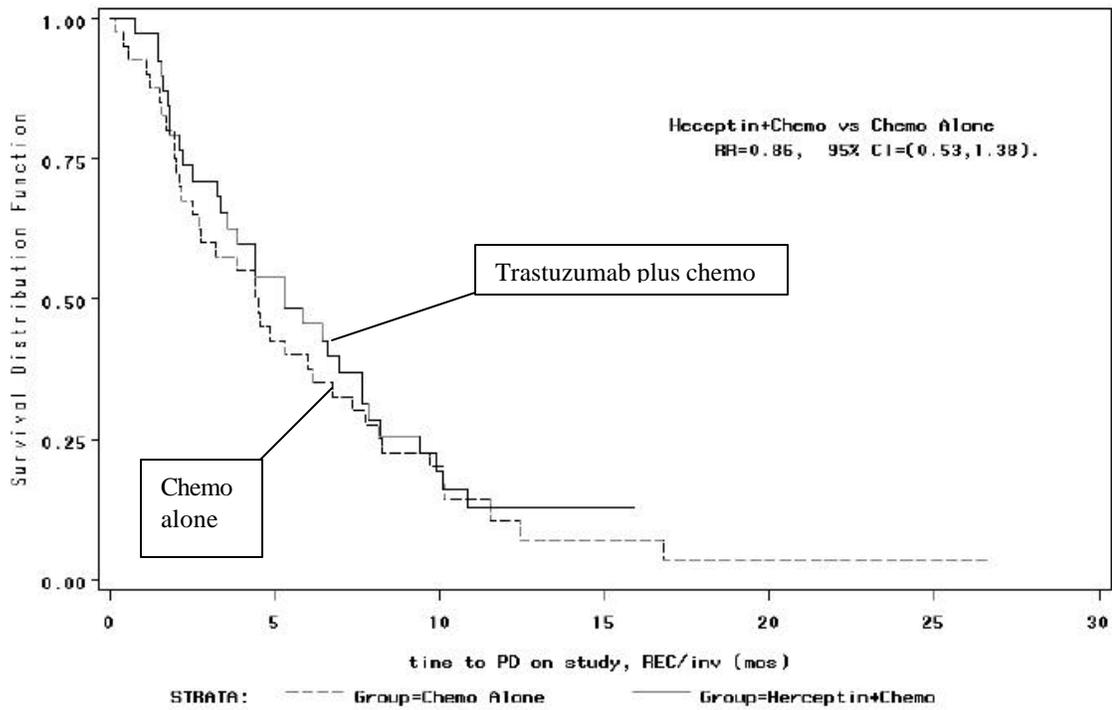


Figure 12. **Time to Progression, FISH-/2+ samples**, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.86 (0.53, 1.38).



Overall Survival

Table 10. Median Overall Survival (months) for H0648g subgroups

	3+ ; H+C	3+; C	2+; H+C	2+ ; C
FISH +	26.3	18.8	21.4	24.5
FISH -	29.1	29.7	19.3	18.4

Figure 13. **Overall Survival, FISH+/3+** samples, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.57 (0.51, 0.89).

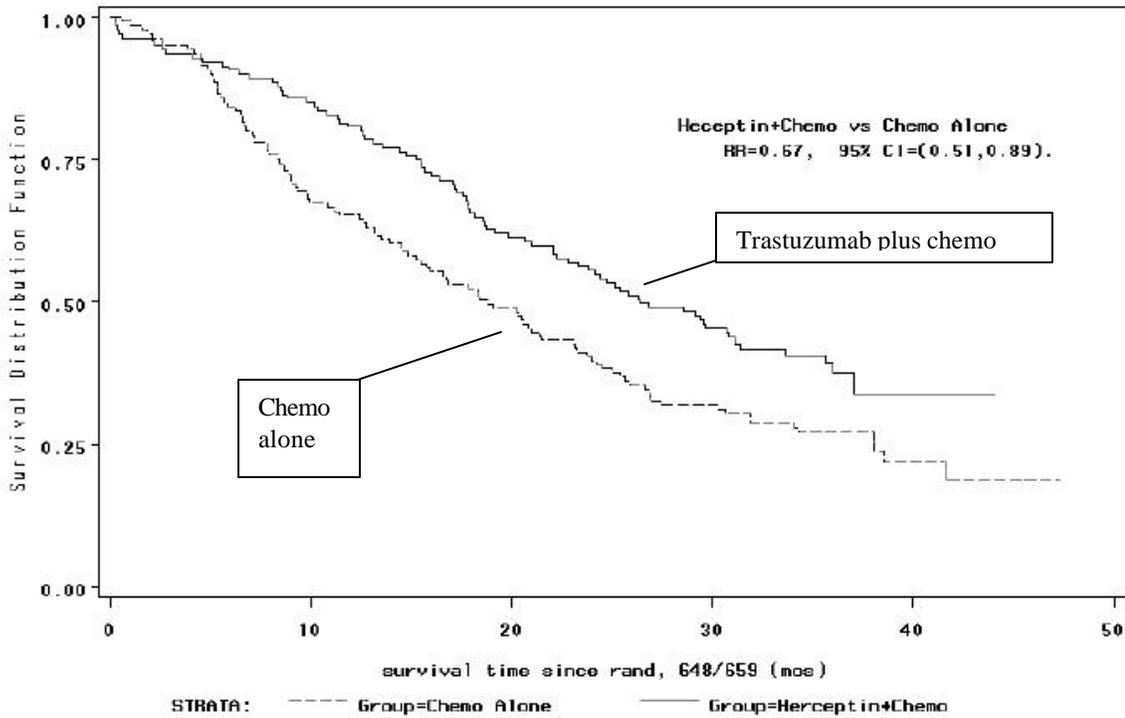


Figure 14. **Overall Survival, FISH-/3+ samples**, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.94 (0.42, 2.11).

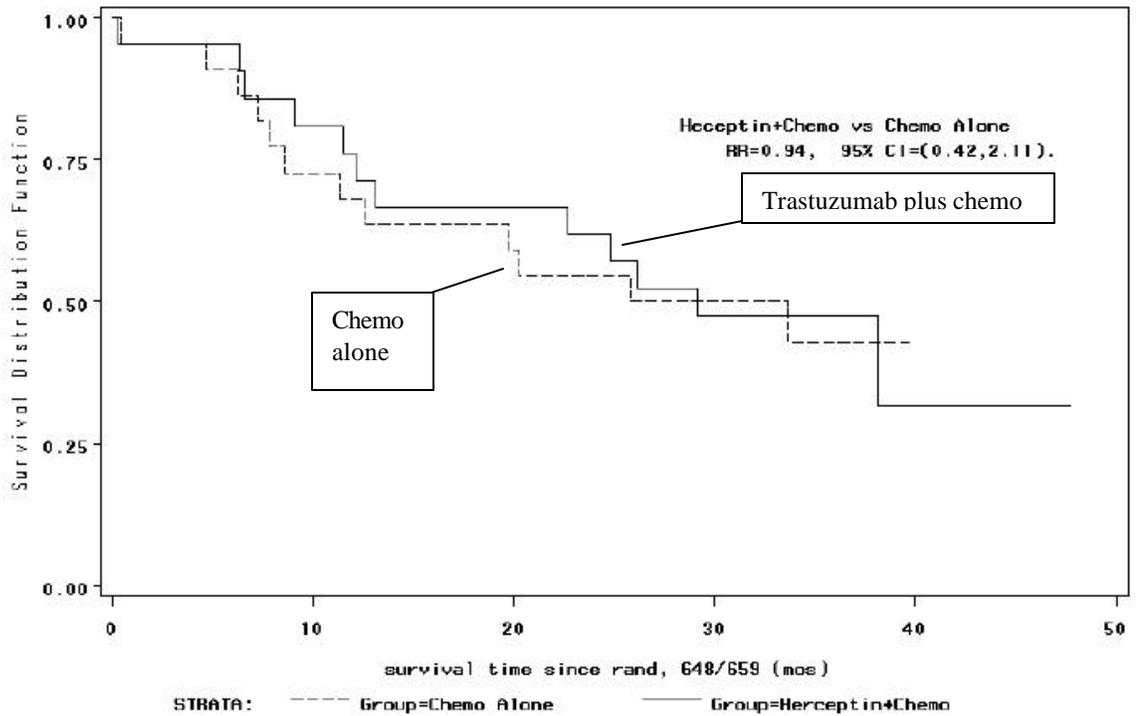


Figure 15. **Overall Survival, FISH+/2+** samples, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 0.98 (0.41, 2.35).

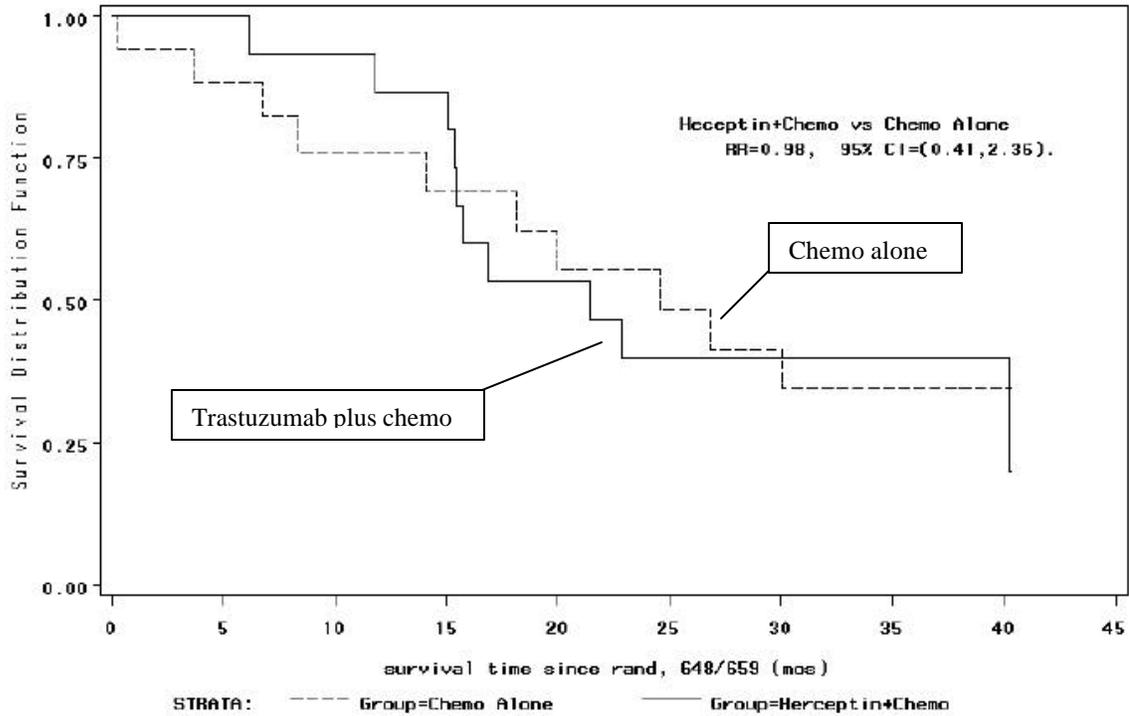
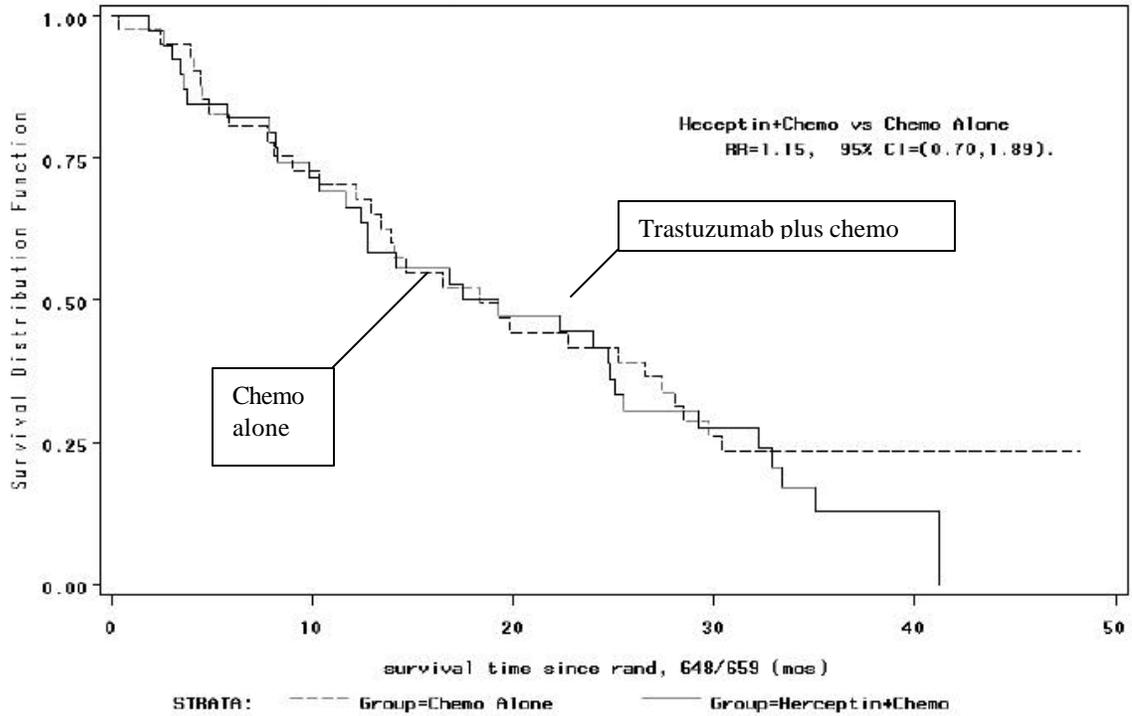


Figure 16. **Overall Survival, FISH-/2+ samples**, for patients enrolled in clinical trial H0648g. Comparison of trastuzumab plus chemotherapy vs chemotherapy alone. The relative risk is 1.15 (0.70, 1.89).



VII. Appendix B
PathVysion Current Package Insert

VIII. Appendix C
HercepTest current package insert

IX. Appendix D
Trastuzumab Package Insert