AD___ 2012_____

Award Number: W81XWH-07-1-0394

TITLE: Utilization of Genomic Signatures to Direct Use of Primary Chemotherapy in Early Stage Breast Cancer

PRINCIPAL INVESTIGATOR: Paul Kelly Marcom, MD

CONTRACTING ORGANIZATION: Duke University, Durham, NC 27708

REPORT DATE: July 2012

TYPE OF REPORT: Final Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE
July 2012

2. REPORT TYPE
Final

3. DATES COVERED
01 July 2007 to 30 June 2012

4. TITLE AND SUBTITLE

Utilization of Genomic Signatures to Direct Use of Primary Chemotherapy in Early Stage Breast Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-07-1-0394

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S) Dr. Paul K. Marcom

E-Mail: marco001@mc.duke.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Duke University
Durham, NC
27708

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The project "Utilization of Genomic Signatures to Direct Use of Primary Chemotherapy in Early Stage Breast Cancer" was an attempt to conduct an integral biomarker validation study of chemosensitivity signatures derived from cell lines for doxorubicin and docetaxel in the context of a standard of care preoperative breast cancer treatment trial using pathologic complete response as a primary endpoint. To execute this trial, a clinical trial support infrastructure was created to be able to obtain full genome expression data in a clinical assay laboratory on breast cancer biopsies, and then interpret that data for relevant predictors. The underlying scientific, preclinical, and clinical work justifying a prospective trial employing these signatures was found to be corrupted, and most of the publications retracted. The clinical trial was therefore terminated. Analyses of available samples using the developed clinical trial infrastructure allowed generation of biologic data from available samples, salvaging potentially useful datasets for investigating breast cancer biology and preliminary validation of signatures from other investigators.

15. SUBJECT TERMS
Breast Cancer, Clinical Trial, Genomics, Biomarkers, Chemotherapy

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(inclde area code)*
U	U	U	UU	31	

Table of Contents

	Page
Introduction…….……..…..	1
Body……..	1
Key Research Accomplishments…….……..	9
Reportable Outcomes……	9
Conclusion……	10
References…….	11
Appendices……	12

Introduction

While begun with great hope and promise, the project "Utilization of Genomic Signatures to Direct Use of Primary Chemotherapy in Early Stage Breast Cancer" encountered problems with data integrity in the basic science underpinning the clinical trial. Consequently, the CDMRP requested early termination of the protocol.. The issues eventuating in this outcome are detailed in a number of sources.[1,2] The findings of the Institute of Medicine were reported in a summary document on March 23, 2012.[3] The clinical trial was permanently closed to accrual, and further payments on the award cancelled.

The clinical trial primary aim was to validation genomic predictors of chemosensitivity based on work conducted by Anil Potti, MD while in the laboratory of Joe Nevins, PhD in Duke's Institute for Genome Sciences and Policy. The reviewers found the promise of chemotherapy predictors to be a strong consideration in funding the grant, particularly in the context of a prospective randomized clinical trial. It is, however, the scientific work underlying these predictors that was found to contain corruption of the original datasets used for validation. While the data corruption detected is the subject of an ongoing scientific misconduct investigation at Duke, the systematic nature of the data changes in this and other work by Dr. Potti point to data manipulation as the only reasonable explanation. The paper describing those signatures has been retracted.[4] Given these developments, continuing with a validation trial for these signatures prospectively was no longer appropriate.

This final report will summarize the work done to institute the clinical trial, including profiling work done on additional samples as part of that work. A considerable genomic dataset was created in this process, and we hope that it will serve some public use and provide some positive outcome from this experience. The project did meet some of the original goals. For example, one goal of the original proposal was to demonstrate the conduct of a prospective preoperative treatment trial using an *integral full-genome expression-based biomarker design* with pathologic complete response as an objective and accepted primary endpoint. While we acknowledge the weaknesses in the effort highlighted by the IOM , we have demonstrated the feasibility of this approach.

Body:

Our original CTRA proposal was motivated by two key concepts: 1) designing a trial with real-time genomic profiling integral to assigning patient treatment; and 2) implementing this in a preoperative trial so that relevant clinical endpoints (pathologic complete response; clinical response) could be assessed. Work in the Duke cancer genomics group had focused on characterizing a number of biologic features using gene expression profiling: 1) prognosis[5]; 2) oncogenic pathway activation[6,7]; and 3) chemosensitivity prediction[8]. Our proposal had the following objectives:

1. Develop and evaluate the clinical infrastructure and methodologies that will be central to the use of genomic profiling to assign patients to treatment.

2. Develop and conduct a Phase II clinical trial that will evaluate the capacity of genomic signatures to improve the efficacy of chemotherapy selection for the individual patient.

3. Make use of the data and results from the Phase II study to develop a follow-up Phase III clinical trial that will compare genomic-guided vs. standard treatment strategies, and will identify novel therapeutic opportunities for those patients likely to be resistant to standard cytotoxic therapies.

Of the "signatures" developed, the chemosensitivity signatures were the most easily integrated into a preoperative treatment protocol that would comport with commonly accepted standards of clinical care. We designed and initiated a prospective Phase II trial integrating the genomic signatures of chemosensitivity to doxorubicin and docetaxel into a preoperative trial.

The following describes the work and accomplishments for each of these objectives:

1. Develop and evaluate the clinical infrastructure and methodologies that will be central to the use of genomic profiling to assign patients to treatment.

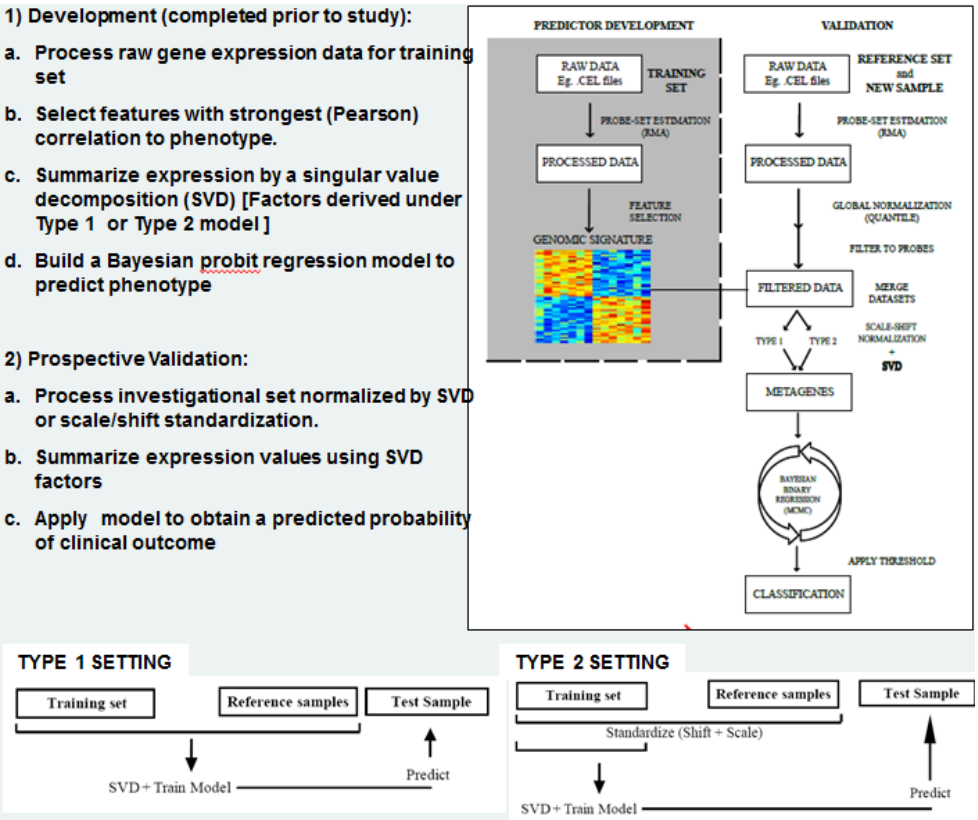


Figure 1: Schematic showing steps for processing expression data from cell line signatures and normalizing one-at-a-time clinical samples into standardization set for generating predictions.

We were able to develop the infrastructure to generate full genome expression data in a CLIA environment (compliant from study start, and approved subsequently), and a bioinformatics pipeline to analyze that data and use it for patient treatment assignments. The chemosensitivity signatures were based on cell-line expression data, so expression data from a comparison set of breast cancers was used to standardize the cell-line based signatures for making predictions of chemosensitivity. This process is outlined in Figure 1 and discussed further in "Intratumor Heterogeneity and Precision of Microarray-Based Predictors of Breast Cancer Biology and Clinical Outcome". [9] This paper also presents the analytic and precision data obtained for the Affymetrix U133A plus 2 arrays and the RNA extraction process that was used in the course of the trial. The chemosensitivity signatures presented in that paper were the versions used in the trial, and had been provided by Dr. Potti as being the signatures used in the prior validation work. Some adjustments were made to transfer the signatures to performance on the U133A plus2 arrays. The precision of these multigene signatures in the replicate dataset is illustrated in Figure 2 from the paper.

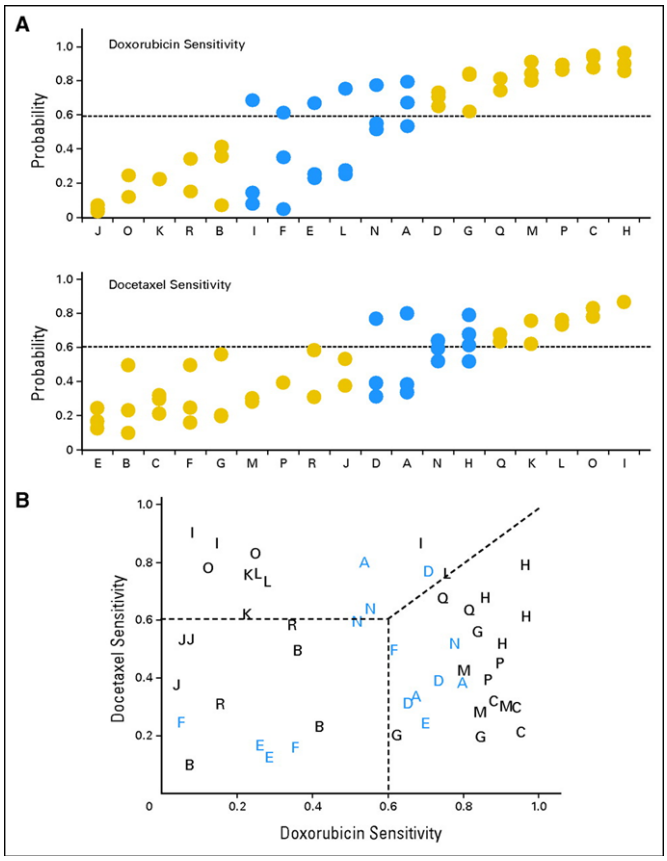


Figure 2: Demonstration of reproducibility of multigene signatures in bioreplicates.

2. Develop and conduct a Phase II clinical trial that will evaluate the capacity of genomic signatures to improve the efficacy of chemotherapy selection for the individual patient.

We also set up a screening process for patient enrollment, detailed in Figure 3.

The biopsies were reviewed by Dr. Joseph Geradts for quality and cellularity, and the best available core used for profiling. The infrastructure for extracting quality RNA was set up in the Clinical Trials Support Facility in collaboration with the IGSP, and was supervised by Dr. Mike Datto. Extracted RNA was then quantitated and then sent to Expression Analysis for hybridization to Affymetrix U133A plus 2 microarrays. The resulting .CEL files were uploaded to a secure server in the IGSP and available only to Dr. William Barry and his biostatistical team. The array data was analyzed by QC metrics to assure adequate signal, and then chemosensitivity predictions using the now discredited signatures were made. The turnaround times and overall success for processing samples are given in Figures 4 and 5.

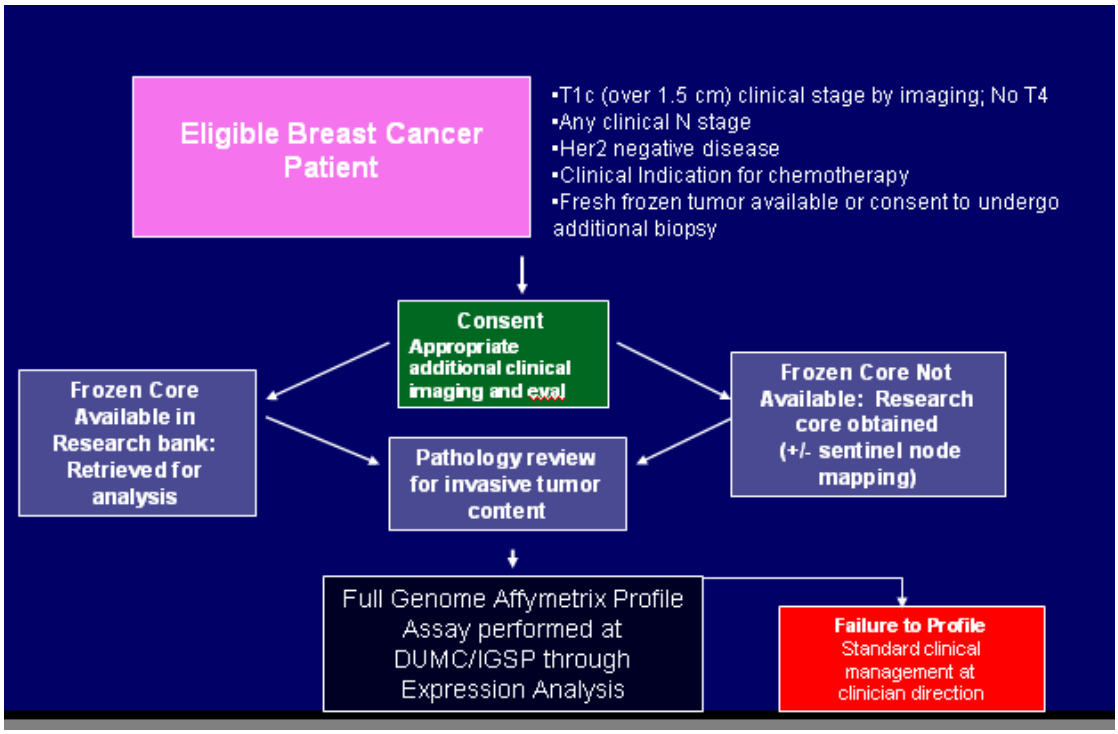
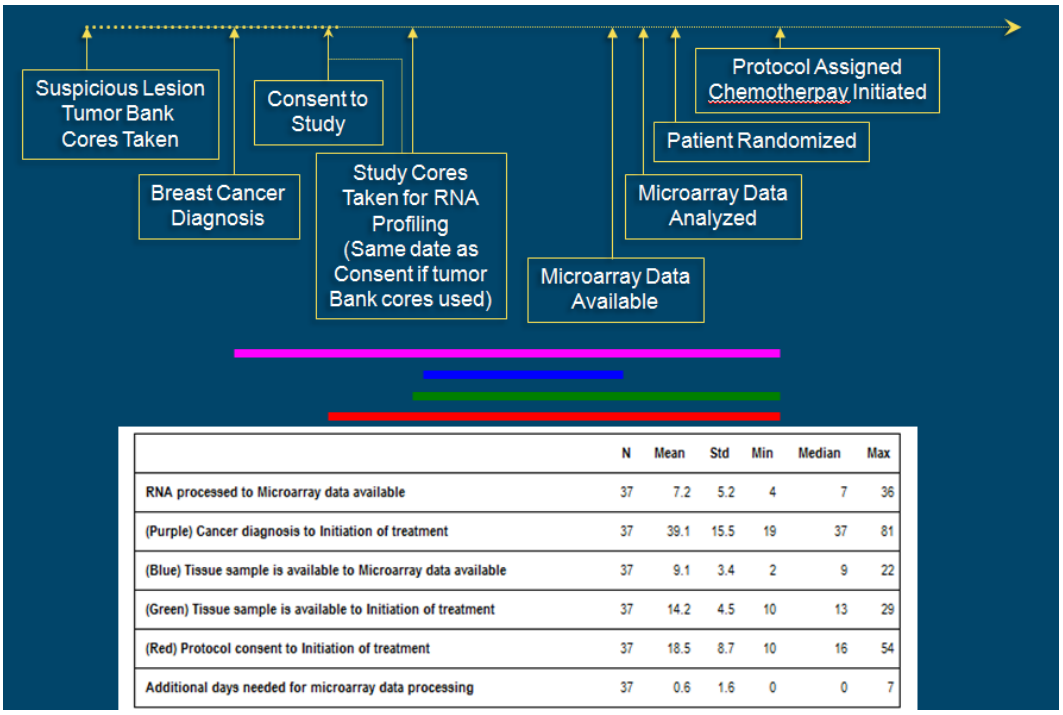


Figure 3: Eligibility criteria, accrual steps, and sample processing in the clinical trial.



	N	Mean	Std	Min	Median	Max
RNA processed to Microarray data available	37	7.2	5.2	4	7	36
(Purple) Cancer diagnosis to Initiation of treatment	37	39.1	15.5	19	37	81
(Blue) Tissue sample is available to Microarray data available	37	9.1	3.4	2	9	22
(Green) Tissue sample is available to Initiation of treatment	37	14.2	4.5	10	13	29
(Red) Protocol consent to Initiation of treatment	37	18.5	8.7	10	16	54
Additional days needed for microarray data processing	37	0.6	1.6	0	0	7

Figure 4: Steps and turnaround times for obtaining, processing, and generating data from core biopsies.

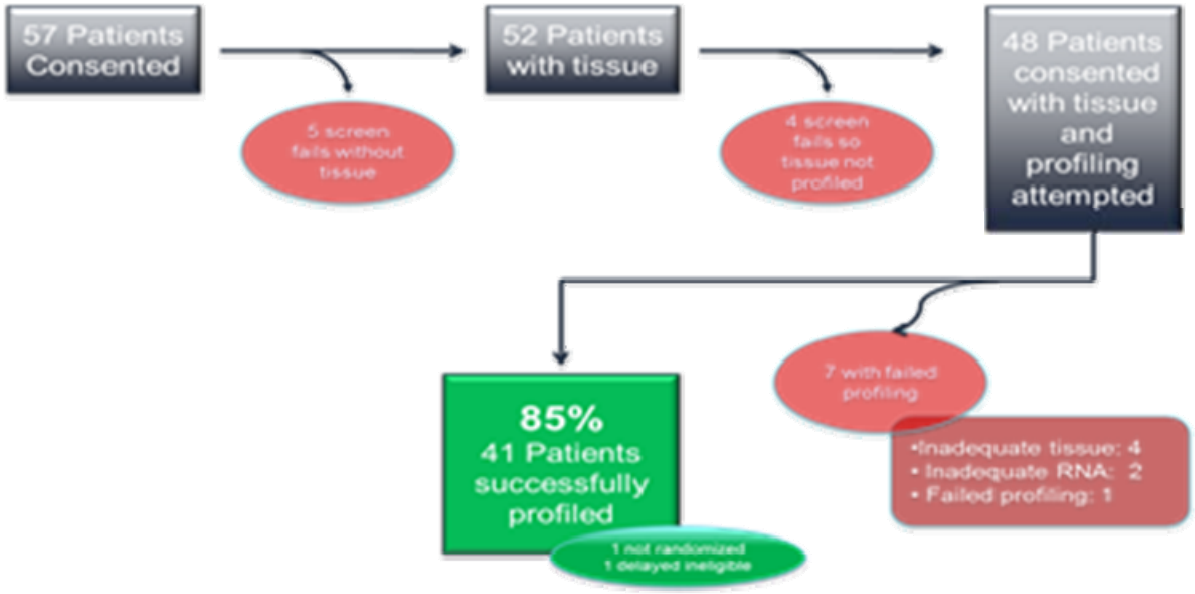


Figure 5: Metrics for sample processing success/failure and obtaining expression data for use in treatment assignments.

	Consented		Randomized and Treated	
	N=57	%	N=39	%
	Age			
Under 40	4	7%	2	5%
40-49	21	37%	14	36%
50-59	17	30%	13	33%
60-69	11	19%	8	21%
70-79	4	7%	2	5%
	Race			
White	41	72%	29	74%
Black	13	23%	7	18%
Hispanic	1	2%	1	3%
Not Reported	2	4%	2	5%
	Tumor Size			
< 4 cm	35	61%	24	62%
≥ 4 cm	21	37%	15	38%
	Histology			
Ductal	45	79%	36	92%
Lobular	6	11%	3	8%
Other	6	11%	0	0%
	ER Status			
ER positive	39	68%	28	72%
ER negative	16	28%	11	28%
ER unknown	2	4%	0	0%

Table 1: Demographic and clinical characteristics of patients.

Table 1 shows the final accrual to the trial before termination. The distribution by race, tumor size, histology, and receptor status is consistent with what was anticipated.

The trial design is given in Figure 6. The design is innovative in being an integral biomarker validation trial. In such a design, obtaining the biomarker (i.e. high quality Affymetrix array data) is required before randomization in the trial. The requirement means that for every patient enrolled in the trial, the required biomarker data will be available. For other designs, where the marker validation is a secondary endpoint, this is not the case, and such designs often fail to meet their objectives because of missing biomarker data.

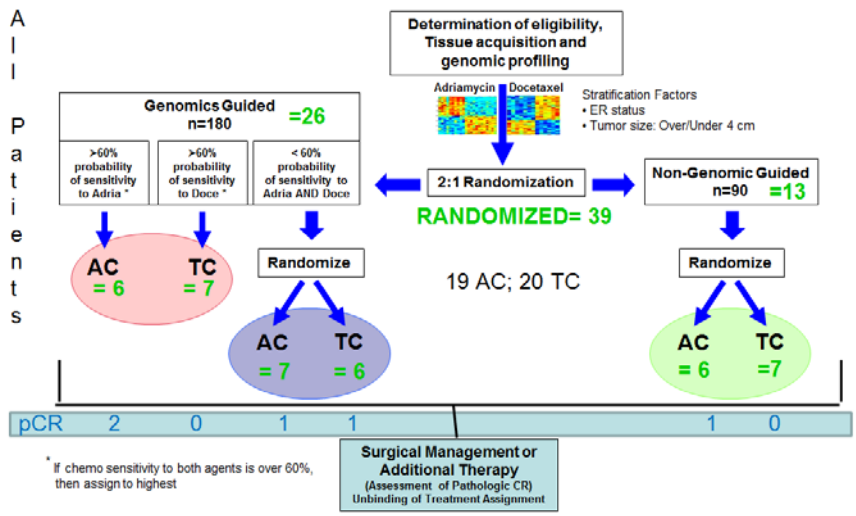


Figure 5: Clinical trial schema and patient assignments, with pCR cases at bottom.

The green numbers in the trial schema show the distribution of patients in the trial. Randomization to genomically-guided versus non-genomically guided arms performed as expected. As well, in the genomically-guided arm, treatment assignments were also uniform among the options, suggesting that the now discredited chemosensitivity signatures were skewed in favor of one regimen or the other.

The primary endpoint for the study was pathologic complete response in the breast, defined as resolution of all invasive cancer. Nodal status was not considered since prechemotherapy sentinel node mapping was allowed in the trial. The treatment assignments were to four cycles of standard dose doxorubicin/cyclophosphamide or docetaxel/cyclophosphamide, both standard regimens for treatment of early stage breast cancer. Patients underwent treatment, and then were clinically reassessed. If additional chemotherapy was thought necessary by the treating physician, that was allowed, but patients were asked to undergo an optional repeat biopsy to established the presence/absence of residual cancer at that time. While that design feature had the potential to undermine the scientific conclusions, it was felt necessary to avoid compromising patient management.

Overall pCR	
13% (95% CI: 4-27%)	
Genomically-guided (n=26)	Control (n=13)
15% (4/26)	8% (1/13)
P=0.64	

AC Treated Cases			
	pCR		
Genomic Prediction	No	Yes	Total
Resistant	9	2	11
Sensitive	6	2	8
Total	15	4	19
Fisher's exact p>0.99			

TC Treated Cases			
	pCR		
Genomic Prediction	No	Yes	Total
Resistant	8	1	9
Sensitive	7	0	7
Total	15	1	16
Fisher's exact p>0.99			

Table 2: Primary analysis of trial endpoint for patients enrolled, and analysis by treatment regimen and prediction of response use discredited chemosensitivity signatures.

The results by pCR in each of the arms and by regimen are given in Table 2. The analysis by regimen was for actual regimen received, not intention to treat, since three of the TC patients did not have definitive surgery following completion of protocol directed therapy and one had an infusion reaction with cycle one and was treated with AC. The trial is, of course, underpowered to show any difference between arms or regimens. However, given the controversy regarding the chemosensitivity signatures, and the supposition that using the signatures to assign treatment might have led to treating them with the wrong drug, we undertook this analysis to see if that appeared to be the case. The results suggest that this is very unlikely.

We will submit the conduct and results of this trial for publication. This will include data that has been presented at several meetings, including ASCO and the ASCO/EORTC/NCI Biomarkers Meeting in October, 2010. [10](See Appendix) Despite the flaws in the underlying science, hopefully the data will prove useful in thinking about the conduct of an integral biomarker-validation trial design. As well, the annotated expression data can provide a potential pilot validation set for biomarkers attempting to distinguish anthracycline and taxane sensitivity in HER2 negative disease.

Table 3: Samples and analyses conducted.

		Sample Set			
Assay/Data		Pre-Op Genomics Trial		Oncotyped Cases	TOTAL Cases (n)
		ER Positive	ER Negative		
Pretreatment	Somatic DNA mutations (Whole Exome Seq Pre AC)	6	6	N/A	12
	Affymetrix Gene Expression Data	29	12	85	126
	Recurrence Score	19	N/A	85	104
	Mammaprint	N/A	N/A	N/A	0
	Circulating blood RNA Expression	29	12	N/A	41
	Proteomics (Pre AC Cases)	9	7	N/A	16
	Baseline MRI	25	10	N/A	16
Post treatment	Somatic DNA mutations (Whole Exome Seq Post AC)	2	4		6
	Affymetrix Gene Expression Data	26	11	N/A	35
	Circulating blood RNA Expression	28	11	N/A	37
	Proteomics (Post AC Cases)	9	6	N/A	15
	Post Treatment MRI	25	10	N/A	15
	Clinical Response	28	11	N/A	39
	Pathologic Response	27	11	N/A	38
	Disease Free Survival	Ongoing for all cases	Ongoing for all cases	Ongoing for all cases	

n= number of potential cases; some with replicates

Data obtained and being analyzed	DNA submitted and sequencing in process	Not applicable or not accomplished

3. Make use of the data and results from the Phase II study to develop a follow-up Phase III clinical trial that will compare genomic-guided vs. standard treatment strategies, and will identify novel therapeutic opportunities for those patients likely to be resistant to standard cytotoxic therapies.

Since the trial was terminated, we could not proceed as planned with this objective. Instead, we sought to salvage the value of the genomic information obtained in this limited dataset by extending the analysis beyond gene expression profiling on the samples in the trial only. The additional samples and analyses conducted were proposed and accepted in the Annual update for Year 4, and are summarized in Table 3. The following analyses are either completed or underway for the data obtained from these samples:

1. **Post Treatment Profiling**
   A total of 37 samples were collected from the post treatment cancers. This set is again unique for examining gene expression changes between doxorubicin and docetaxel treated patients. While this dataset will be small, that is true of most similar datasets; they are challenging to develop. Nevertheless, this sample set can be used for discovery purposes and as a potential validation set for predictors [11]

2. **Peripheral Blood Gene Expression Profiling**
   PAX gene tubes for isolation of peripheral circulating RNA were obtained at baseline and after protocol directed therapy on all patients. The RNA has been extracted from these samples, and will be profiled using Affymetrix arrays to look and changes in circulating RNA. These changes can be compared to changes in the primary tumor and provide a pilot set for development of response and toxicity prediction signatures.

3. **Expansion of bioreplicate cohort**
   Most of the cases enrolled in the trial had at least one additional pretreatment core taken. We have profiled those cores to expand the replicate dataset already published. This work was recently presented at the 2012 San Antonio Breast Cancer Symposium, and will be expanded further to look at reproducibility as a function of tumor cellularity and to look at reproducibility of published prognostic and predictive signatures, and is included in the appendix.

4. **Pre/Post Breast MRI Patterns and Gene Expression Profiles**
   This dataset can also be analyzed for baseline breast MRI patterns and gene expression profiles. While some work has been done in this area, the tumors are generally larger, locally advanced cancers.[12] This dataset has smaller tumors represented. As well, the data on focused regimens is unique. Prior work, including some at Duke, suggests that MRI patterns of enhancement can predict response. We can look at similar imaging patterns in this cohort of patients, and also analyze by subtype.

5. **Cross-platform Genomic Prognostic Signatures.**
   Using the 19 patients in the study with clinical Oncotype Recurrence Scores as a starting set, we have assembled a cohort of 104 cases from the SPORE tumor registry with concurrent Oncotype scores as a comparison set. This analysis was presented in poster form at the 2012 San Antonio Breast Cancer Symposium and is included in the appendix.

6. **Proteomic Analysis Somatic Gene Mutations**
   We have done proteomic profiling of pre/post AC treatment cores on 17 patients using 2DLC-MS/MS analysis. Initial analysis of the data has been performed and indicates that this approach is possible, and that changes in protein expression can be detected. We are attempting to integrate this data with gene expression data and complete the primary analysis for protein expression changes that might be associated with anthracycline sensitivity and resistance.

7. **Somatic Gene Mutations**
   With the processing of replicate cores, we were able to extract tumor DNA from 12 pre-treatment cores and 6 matched post-treatment cores from the AC treated cohort of patients. In a separate collaboration, these samples are being analyzed with whole exome sequencing to look at potential changes is somatic mutations detected. As well, four of these cases had pathologic complete responses, so a preliminary discovery effort can be made to look for mutations predicting AC response.

8. **Lymph Node Predictor**
   Previous work in the Duke Genomics group focused on development of a predictor of lymph node involvement.[5,13,14] The clinical applicability of the lymph node predictor was unclear at the time this work was performed. However, with new studies suggesting that completion axillary dissection is not necessary for patients with limited lymph node involvement, the performance characteristics needed for a lymph node predictor have changed. The datasets described here provide a further opportunity to validate and/or further develop a lymph node predictor.

Key Research Accomplishments

- Demonstration of the performance of the clinical trial infrastructure, including rates of successful profiling, quality assurance of genomic data, and time metrics for assay performance.
- Randomization of 39 patients in the prospective trial, with associated biospecimens (including frozen tumor, germline DNA, and blood).
- Generation of a small but valuable dataset for potential validation of published chemosensitivity signatures from other groups, particularly for comparing anthracycline- and docetaxel-based treatment in isolation.
- Dataset of expression profiling analysis of 104 estrogen receptor positive early stage breast cancers annotated with clinical Oncotype Recurrence Scores.
- Dataset of expression profiling analysis of 46 cases with between 2 and 5 bioreplicate cores of varying cellularity useful for studying reproducibility of gene expression based prognostic and predictive signatures.
- Proteomic data on 17 matched pre/post AC treatment cores.
- Peripheral/blood gene expression profiles on 34 patients pre/post chemotherapy exposure.
- To be completed with support from other funds: whole exome sequencing on pre/post AC treated cases.

Reportable Outcomes

The following work related to this project has been presented at meetings and/or published. Please see the documents appended at the end of this report.

- Poster 1: "Development and Implementation of Genomic Predictors of Chemotherapy Response for Guiding Preoperative Therapy in a Prospective Breast Cancer Trial", NCI Translational Science Meeting, November 5-7, 2009.
- Poster 2: "A randomized phase II trial evaluating the performance of genomic expression profiles to direct the use of preoperative chemotherapy for early-stage breast cancer" ASCO Annual Meeting, 2010.
- Poster 3: "Utilization of Genomic Signatures for Chemotherapy Response in Prospective Clinical Studies", ASCO Annual Meeting, 2010. (Poster Discussion link on ASCO Virtual Meeting is given in Appendix)
- Article 1: "Intratumor Heterogeneity and Precision of Microarray-Based Predictors of Breast Cancer Biology and Clinical Outcome", Journal of Clinical Oncology, 29 (13): 2198, May 1, 2010.
- Poster 4: "Pathologic response analysis for a genomically guided preoperative chemotherapy trial", Era of Hope Meeting, 2011.
- Poster 5: "Generation of real-time full-genome expression data for treatment assignment in a prospective breast cancer trial", Era of Hope Meeting, 2011.
- Poster 6: "A Cross-Platform Comparison of Genomic Signatures and OncotypeDx Score: Discovery of Potential Prognostic/Predictive Genes and Pathways", SABCS, 2012.
- Poster 7:" Retrospective evaluation of precision of gene-expression-based signatures of prognosis and tumor biology in replicate surgical biospecimen from patients with breast cancer", SABCS, 2012.

Conclusion:

Given the collapse of the scientific rationale for the clinical trial, we have attempted to salvage as much meaningful data from the project as possible. We maintain that patients still received optimal standard of care treatment, in a study design that allowed individualization of treatment while also meeting the scientific goals. All of the patients were made aware of the scientific misconduct issues. The clinical team undertook this validation trial in good faith, believing that the peer-reviewed published science justified the prospective study, and that the trial design abided by standard of care treatment. Importantly, as outlined above, the trial also held promise for establishing the process for integrating a complex biomarker into clinical trial design, allowing for correlation of the underlying biology with other clinically relevant factors. Hopefully, we can still in some way honor the patients who agree to participate in this study by the work detailed above.

In retrospect, the data supporting the conduct of the trial was not adequate to justify proceeding. This report has not attempted to address all of the issues raised in the IOM report or the controversy related to the scientific misconduct. Unfortunately, it is undoubtedly true that the ensuing scandal has damaged the field of translational genomics. While those issues went well beyond this project, we profoundly regret the damage done and any connection to this project. Duke University has investigated this matter and developed policies to ensure the integrity of data supporting translational clinical research.[15,16] While we clearly fell short of the goal, we

hope the data generated will provide helpful clinic-genomic datasets to the breast cancer research community, and will release the data generated in a manner consistent with the Translational Medicine Quality Framework policies.

References:

1. The Cancer Letter. (Accessed at http://www.cancerletter.com/.)
2. Baggerly K, Coombes, K. . Deriving chemosensitivity from cell lines: Forensic bioinformatics and reproducible research in high-throughput biology. Annals of Applied Statistics 2009;3:1309-34.
3. Trials CotRoO-BTfPPOiC, Services BoHC, Policy BoHS, Medicine Io. Evolution of Translational Omics: Lessons Learned and the Path Forward: The National Academies Press; 2012.
4. Potti A, Dressman HK, Bild A, et al. Retraction: Genomic signatures to guide the use of chemotherapeutics. Nat Med 2011;17:135-.
5. Huang E, Cheng SH, Dressman H, et al. Gene expression predictors of breast cancer outcomes.[comment]. Lancet 2003;361:1590-6.
6. Huang E, Ishida S, Pittman J, et al. Gene expression phenotypic models that predict the activity of oncogenic pathways. Nature Genetics 2003;34:226-30.
7. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 2006;439:353-7.
8. Potti A, Dressman HK, Bild A, et al. Genomic signatures to guide the use of chemotherapeutics. Nat Med 2006;12:1294-300.
9. Barry WT, Kernagis DN, Dressman HK, et al. Intratumor Heterogeneity and Precision of Microarray-Based Predictors of Breast Cancer Biology and Clinical Outcome. Journal of Clinical Oncology 2010;28:2198-206.
10. Marcom PK, Datto MB, Barry WT, et al. Implementation of genomic predictors of chemotherapy response for guiding preoperative therapy in a prospective breast cancer trial. In; 2009:11057-.
11. Soo-Chin Lee XX, Wee-Joo Chng, Mark Watson, Yi-Wan Lim,, Chiung-Ing Wong PI, Norita Sukri, Siew-Eng Lim, Hui-Ling Yap,, Shaik Ahmad Buhari PT, Jiayi Guo, Benjamin Chuah,, Goh HLMaB-C. Post-treatment tumor gene expression signatures are more predictive of treatment outcomes than baseline signatures in breast cancer. Pharmacogenetics and Genomics 2009;19:833-42.
12. Esserman L, Kaplan E, Partridge S, et al. MRI Phenotype Is Associated With Response to Doxorubicin and Cyclophosphamide Neoadjuvant Chemotherapy in Stage III Breast Cancer. Annals of Surgical Oncology 2001;8:549-59.
13. Pittman JL, Tebbit CL, Black EP, et al. Predictive models that combine multiple forms of genomic and clinical data to achieve personalized prediction of outcomes in breast cancer. Breast Cancer Research and Treatment 2004;88:S21-S2.
14. Pittman J, Huang E, Dressman H, et al. Integrated Modeling of Clinical and Gene Expression Information for Personalized Prediction of Disease Outcomes. Proceedings of the National Academy of Sciences of the United States of America 2004;101:8431-6.
15. Califf RM, Kornbluth S. Establishing a Framework for Improving the Quality of Clinical and Translational Research. Journal of Clinical Oncology 2012;30:1725-6.

16. Duke Medicine Translational Medicine Quality Framework Committee(5 1, 2011) A framework for the quality of translational medicine with a focus on human genomic studies: Principles from the Duke Medicine Translational Medicine Quality Framework Committee. (Accessed at http://medschool.duke.edu/research/translational-medicine-quality-framework.)

Appendices

1. Funded personnel during conduct of the project.
2. Paper and Abstracts cited in Reportable Outcomes.

Funded Personnel:

Paul Kelly Marcom MD- Overall study Principal Investigator, years 1-5
Anil Potti MD- Scientific advisor, years 1-3. No role in trial conduct or analysis of any data in this report.
John Olson MD, PhD- Surgical co-investigator
Joe Nevins PhD- Scientific advisor, years 1-5
Joseph Geradts MD- Study breast pathologist
Sujata Ghate MD- Study breast radiologist
Bill Barry PhD- Study biostatistician
Jeff Marks PhD- Scientific investigator
Mike Datto MD, PhD- Study molecular pathologist
Bob Annechiarico- Study Bioinformatics
Traci Foster RN, BSN- Overall project coordinator years 1-4
Jamie Cuticchia PhD- bioinformatics
Geoff Ginsburg MD, PhD- IGSP investigator and Director of Genomic Trials
Tim Veldman- Project manager for IGSP
Ryan Griffis- Molecular lab technician
Jvonne Hunter- Molecular lab technician
Kirk Gray- Cancer Center bioinformatics and Oracle case report form developer
Holly Dressman- IGSP microarray facility director
Caroline Lentz- Clinical protocol nurse
Vlayka Liotcheva- Clinical Research Coordinator, years 2—3, project coordinator years 4-5
Pam Isner- Molecular lab technician
April Coan- biostatistician
Pankaj Agarwal- biostatistician
Laurie Rouse- assistant project coordinator

Development and Implementation of Genomic Predictors of Chemotherapy Response for Guiding Preoperative Therapy in a Prospective Breast Cancer Trial

P. Kelly Marcom, MD, William Barry, Ph.D.,Michael B Datto, MD, PhD, Joseph Geradts, MD, Traci L Foster, RN, Laurie Rouse, Holly K Dressman, PhD,
John Olson Jr., M.D., Ph.D., Geoff Ginsburg, MD, PhD, Anil Potti, MD,
H. Kim Lyerly, MD, Joseph R Nevins, PhD

Duke Multidisciplinary Breast Program

Email: marco001@mc.duke.edu

DUKE INSTITUTE FOR GENOME SCIENCES & POLICY

Duke Comprehensive Cancer Center
DUKE UNIVERSITY HEALTH SYSTEM

ABSTRACT

Background: Genomic assays have been shown to provide the potential for personalized approaches to breast cancer therapy. While assays based on fixed tissues offer greater convenience, the use of fresh tissue samples provides a much broader opportunity to interrogate the critical underlying biology. Full-transcriptome assays using microarrays are challenging, but have the advantage of providing multiple prognostic and predictive signatures in one assay. Using an Affymetrix platform, we have developed genomic predictive signatures for various clinically relevant cancer characteristics. Previous credentialing studies have demonstrated the sensitivity and specificity of these signatures in retrospective studies. We report here on the investigation of the analytic performance of these signatures in biologic replicates, and their application in a prospective randomized clinical trial designed to validate predictors of chemosensitivity. Methods: To investigate the impact of intratumoral heterogeneity on the genomic signatures, patients with multiple frozen cores were identified in the Duke Breast SPORE tissue repository. Cores were assessed for percent invasive cancer cellularity, primary tumor size, and standard biomarker assessments. RNA was hybridized to Affymetrix H133 Plus 2.0 microarrays and the data was used to evaluate gene expression signatures generated according to previously identified predictors of sensitivity to adriamycin and docetaxel (Potti et al, Nat Med, 2006). Genomic predictors of hormone receptor status were applied to post- processed array data, and compared to single-patient measures from IHC/FISH. The resulting analyses support the currently enrolling trial, "Performance of Genomic Expression Profiles to Direct the Use of Preoperative Chemotherapy for Early Stage Breast Cancer" a prospective randomized trial designed to validate genomic signatures for predicting response to doxorubicin or docetaxel treatment in HER2 negative cancers.

Results: Fifty-one individual samples from 18 patients were profiled to investigate the contribution of intratumoral heterogeneity to signature variation. The interclass correlation for doxorubicin and docetaxel sensitivity predictor in the replicate samples was 0.72 and 0.64 respectively ($p < 0.0001$). Expression data was also analyzed for a novel predictor of ER pathway activation. Predicted ER status among replicates showed perfect concordance using this pathway signature. The infrastructure has been established for conducting microarray analysis in support of the clinical trial, and provides microarray data with an approximately 5 day turnaround from biopsy to data availability. Multi-center accrual to the study is underway. Conclusions: Full genome microarray expression profiles are robust and reproducible, and can be practically obtained and applied in the context of a prospective clinical trial. This single assay can provide data that can be analyzed for a variety of clinically useful prognostic and predictive signatures.

METHODS:

In this prospective trial, genomic assessment is performed on tumor samples in real-time using Affymetrix U133Plus2.0 microarrays. Previously described "signatures" are then used in conduct of the clinical trial. Figures one and two outline the procedures for tissue acquisition and the design of the clinical protocol, respectively. The results describe the metrics for performing this assay in the context of this trial.

Figure 1: Patient eligibility and tissue acquisition

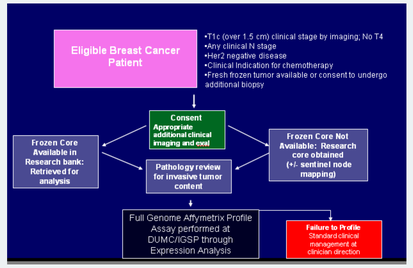


Figure 2: A Randomized Phase II Trial Evaluating Performance of Genomic Expression Profiles to Direct the Use of Preoperative Chemotherapy for Early Stage Breast Cancer

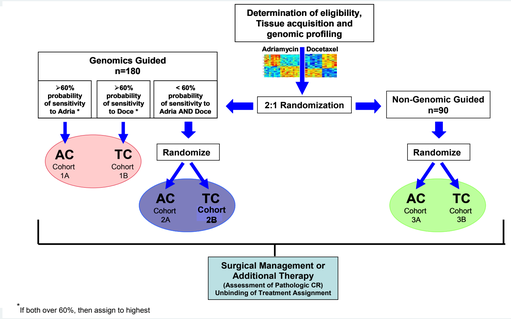
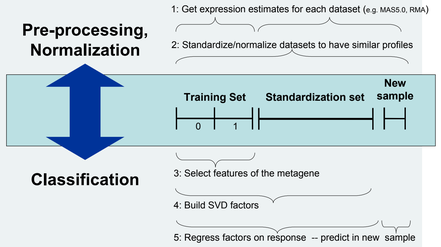
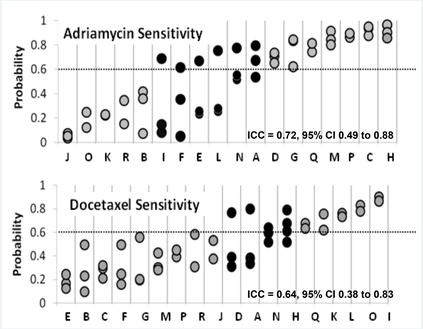


Figure 3: General steps in prospective classification of microarray samples



RESULTS:

Figure 4: Signature Concordance on Bioreplicate Cancer Samples



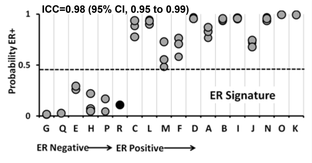


Table 1: Sample collection and microarray profiling success

REASON	N	%
Screening Failures		
Histology Q/C Failure	1	3
RNA Q/C Failure	1	3
Array Q/C Failure	0	0
Technical or Processing Error	1	3
Not Eligible	0	0
Withdrew Consent	3	10
Total Screening Failures	**6**	**21**
Randomized	23	79
Total	**29**	**100**

•Patients 1004 and 1018 declined second biopsies and withdrew consent

•Patient 1023 had a low Oncotype score and withdrew consent to be treated with chemotherapy.

•Patient 1007 had no invasive tumor in the research cores and was classified as a Histology Q/C failure.

•Patient 1013 had multi-centric tumor, but did not have enough RNA for genetic testing. They were classified as an RNA Q/C failure.

•Patient 1015 was a RNA hybridization failure. They were classified as a technical or processing error. (See Figure X for QC plot of principal components for this analysis).

88% successful profiling rate
(95% CI: 70% - 98%)

Excludes three patients who withdrew consent during screening

Figure 5: Microarray quality control (QC) plot for sample 1015 identified hybridization failure

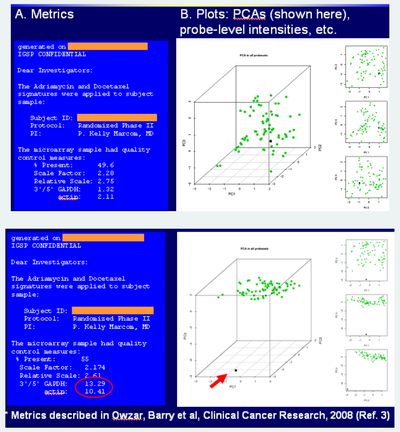
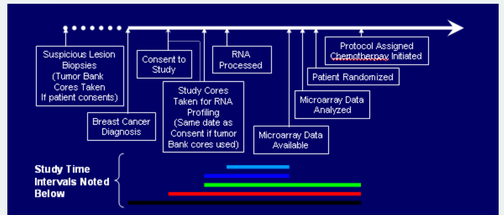
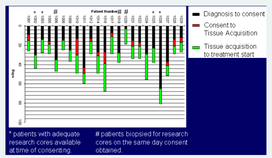


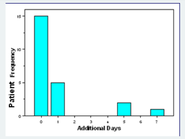
Figure 6: Timeline for protocol steps



Study Time Interval (Days)	N	Mean	SD	Minimum	Median	Maximum
(Lt Blue) RNA processed to micro-array data available	23	6.6	1.9	4	7	9
(Dark Blue) Sample is available to microarray data Available	23	9.8	3.7	5	9	22
(Green) Sample is available to Initial of Treatment	23	14.3	5.2	8	13	29
(Red) Protocol Consent to Initiation of Treatment	23	18.0	7.6	10	16	41
(Black) Cancer Diagnosis to Initiation of Treatment	23	40.3	15.4	19	37	81

Figure 7: Individual patient timelines and impact of profiling on treatment initiation





Eight cases required one (5) or more (3) additional days for microarray data for making treatment assignment. (Mean= 2.75 days)

CONCLUSIONS:

1. True biologic replicates demonstrate substantial concordance for complex multi-gene expression signatures.
2. Real-time, one-at-a-time, full genome microarray expression analysis using fresh frozen tumor samples and an Affymetrix platform as a clinical assay is feasible.
3. Profiling success rates and turn-around times are acceptable for using this approach for making genomically guided treatment assignments in a preoperative therapy breast cancer trial.
4. In addition to the chemosensitivity predictions being evaluated in the current trial, the same full genome expression data can be analyzed for other clinically relevant prognostic and predictive factors.

Supported by Duke Breast SPORE grant P50CA068438, CDMRP Department of Defense Clinical Trial award grant W81XWH-07-1-0394, a grant from the V Foundation.

A randomized phase II trial evaluating the performance of genomic expression profiles to direct the use of preoperative chemotherapy for early-stage breast cancer

P. K. Marcom[4,5], W. T. Barry[1,5], M. B. Datto2,5, J. Geradts[2], J. Olson Jr.[3,5], J.. Marks[3,5], G. Lyman[4], G. Ginsburg[5] ,A. Potti[4,5], J. R. Nevins[5]

Duke Multidisciplinary Breast Program

1) Dept. of Biostatistics and Bioinformatics, 2) Dept. of Pathology, 3) Dept. of Surgery, 4) Dept. of Oncology

Email: marco001@mc.duke.edu

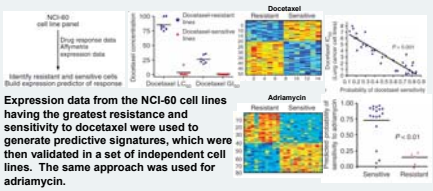
ABSTRACT

Background: Genomic assays have been shown to provide the potential for personalized approaches to breast cancer therapy. Full-transcriptome assays using microarrays are challenging, but have the advantage of providing multiple prognostic and predictive signatures in one assay. We have developed genomic predictive signatures for various clinically relevant cancer characteristics, including sensitivity to cytotoxic chemotherapies. We have investigated the analytic performance of these signatures in biologic replicates, and begun applying them in this prospective randomized clinical trial. Methods: Eligible patients have intact HER2 negative primary breast cancers with tumor size over 1.5 cm and any clinical lymph node status. A fresh frozen sample of the primary tumor must be obtained. After pathology assessment, RNA is analyzed on an Affymetrix H133 Plus 2.0 microarray. The expression data is used to predict doxorubicin (A) and docetaxel (T) sensitivity using previously defined signatures and analytic techniques developed to make one-at-a-time predictions from microarray data. Patients are then assigned to four cycles of standard doxorubicin/cyclophosphamide (AC) or docetaxel/cyclophosphamide (TC) chemotherapy in a blinded process randomizing in a 1:2 ratio to a control arm (random assignment to either AC or TC) or a genomically guided (GG) chemotherapy arm. In the GG arm, patients are assigned to AC or TC treatment if the predicted sensitivity is over 0.6 for A or T, respectively. If predictions for both agents are over 0.6, the higher score is used; if both scores are below 0.6, patients are randomly assigned to AC or TC. The primary endpoint is pathologic complete response (pCR) in the breast, with the primary objectives: 1) determine whether genomic profiling for drug-sensitivity can improve the pCR rate as compared to randomly assigned therapy; and 2) determine whether genomic profiling can identify drug-sensitive and drug-resistant patients including a comparison of subgroups for the two individual regimens. The accrual goal is 270 patients randomized; currently, 42 patients have been consented and 32 randomized. (Current accrual updated in poster)

BACKGROUND:

Genomic signatures can provide an array of predictive biomarkers to guide selection of therapy. Using in vitro approaches can potentially accelerate this discovery process and expand the array of predictive tools. We have described *in vitro* cell culture experiments used to develop predictive genomic signatures, including commonly used cytotoxic chemotherapeutics (Refs 1-3). Figure One outlines the approach for docetaxel (T) and adriamycin (A).

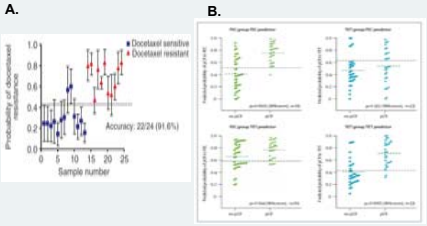
Figure 1: *in vitro* modeling



Expression data from the NCI-60 cell lines having the greatest resistance and sensitivity to docetaxel were used to generate predictive signatures, which were then validated in a set of independent cell lines. The same approach was used for adriamycin.

Figure 2: Retrospective Clinical Validation

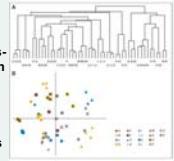
These signatures were then applied to a number of publically available clinical datasets with available genomic expression data to assess their performance for predicting sensitivity and resistance.



In a preoperative breast cancer trial by Chang et al, (Ref. 4) patients treated with single agent docetaxel were assessed for clinical response, and genomic profiling data obtained on the cancers. The in vitro signature demonstrated good accuracy for predicting sensitive/resistant tumors (A). In a second preoperative trial, Bonnefoi et al genomically profiled cancers treated with standard Fluorouracil/Epirubicin/Cyclophosphamide (FEC) or Epirubicin/Taxotere followed by Taxotere (TET). A combination of *in vitro* derived signatures predicted clinical response (B).(Ref.5)

Figure 3: Microarray Signature Precision

To develop the use of microarray data as a clinical assay that allows simultaneous assessment of multiple predictive signatures, including those for A and T, we profiled replicate samples from the Duke Breast SPORE, performed hierarchical clustering (A) and principal component analysis (B) to assess concordance of genomic profiling. The signatures for A and T were also assessed (C) and showed intraclass correlation coefficients (ICC) of 0.72 and 0.64, respectively. These data provided evidence of reproducible precision for these and other genomic signatures. (Ref. 7) Full gene lists for the signatures are provided in this reference.



C

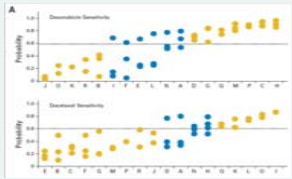
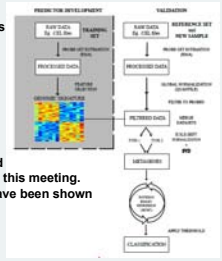


Figure 4: Prospective Use of Gene Signatures

A key aspect of employing *in vitro* derived gene signatures prospectively requires normalization of the cell line training set data into a standardization set of tumor expression data. This allows one-at-a-time prediction using the genomic signature of interest. The statistical methods developed for this approach are described further in abstract #10513 from this meeting. (Ref. 8). Similar approaches have been shown to work by others (Ref. 9).



TRIAL DESIGN

Based on this work, we are conducting a prospective random-ized preoperative clinical trial assessing the use of the A and T signatures to choose between two standard of care regimens. In this design, the biomarker is integral to the trial design. The eligibility criteria, process for obtaining tissue, and clinical schema are shown below. The method of treatment assignment to AC or TC is double blinded.

Figure 5: Patient eligibility and tissue acquisition

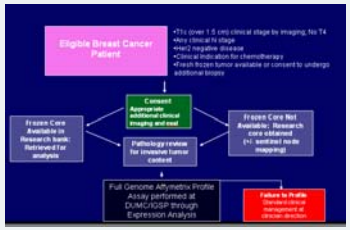
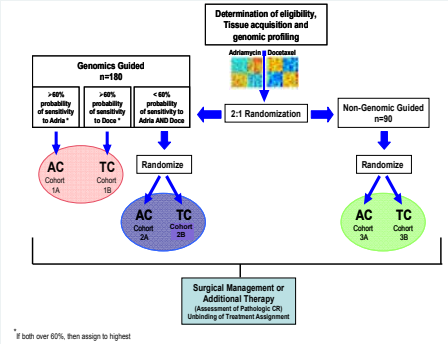


Figure 6: Trial Schema



Primary Objectives and Statistical Design:

1. To determine in early stage breast cancer treated with PST whether genomic profiling for drug-sensitivity can improve the pCR rate as compared to random assignment of patients to therapy.

2. To determine in early stage breast cancer treated with PST whether genomic profiling can identify drug-sensitive and drug-resistant patients including a comparison of subgroups for the two individual regimens (i.e. AC and TC).

Primary objective 1 is to test for a difference between arms in pCR rate using a one-sided Type 1 error rate of 0.05. The null hypothesis is that the pCR rate in both arms is 0.13. Given that the true pCR rates are 0.13 versus 0.26, the chi-square test of proportions has 82% power of rejecting the null hypothesis. Given that the true pCR rates are 0.13 versus 0.24, the chi-square test has 72% power. Even if the null hypothesis is not rejected, Primary Objective 2 will have reasonable power to examine the difference in pCR rates between predicted drug sensitive and predicted drug resistant for the individual regimens. An interim analysis is scheduled at 90 patients to assess for futility.

CURRENT STATUS

- 50 patients have been consented at DUMC
- 38 patients have been randomized
- 4 additional clinical sites have recently been approved for enrollment.

REFERENCES

1. Potti A. et al: Genomic signatures to guide the use of chemotherapeutics. Nat Med. 12:1294-1300, 2006
2. Bild AH et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 439:353-7, 2006
3. West M et al: Predicting the clinical status of human breast cancer by using gene expression profiles. PNAS 98: 11462-7, 2001
4. Chang J et al: Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet, 2003
5. Bonnefoi et al: Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. Lancet Oncology, 2007.
6. Marcom et al: Implementation of genomic predictors of chemotherapy response for guiding preoperative therapy in a prospective breast cancer trial, ASCO Meeting Abstracts May 20 2009: 11057.
7. Barry WT et al: Intra-tumor heterogeneity and precision of microarray-based predictors of breast cancer biology and clinical outcome. J Clin Oncol. 28(13), 2198-206, 2010
8. Barry WT et al: Utilization of Genomic Signatures for Chemotherapy Response in Prospective Clinical Studies.Abstract #10513 ASCO. 2010.
9. Lee JK et al: Prospective comparison of clinical and genomic multivariate predictors of response to neoadjuvant chemotherapy in breast cancer. Clin Can Res 2010; 16: 711.

Supported by CDMRP Department of Defense Clinical Trial award grant W81XWH-07-1-0394, Duke Breast SPORE grant P50CA068438, and a grant from the V Foundation.

Utilization of Genomic Signatures for Chemotherapy Response in Prospective Clinical Studies

William T. Barry[1,5], Chaitanya R. Acharya[5], Michael B. Datto[2,5], Holly K. Dressman[5], P. Kelly Marcom[3,5], Neal Ready[3,5], Geoff S. Ginsburg[3,5], Anil Potti[3,5], Joseph R. Nevins[5]

1) Dept. of Biostatistics and Bioinformatics, 2) Dept. of Pathology, 3) Dept. of Medicine
5) Institute for Genome Science and Policy, Duke University, Durham NC.

Email: bill.barry@duke.edu

Abstract

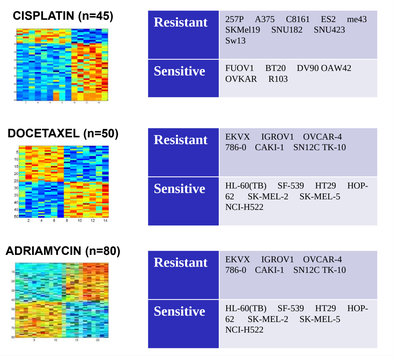
Background: The development of genomic signatures as predictive biomarkers of chemotherapeutics provides the potential to guide therapeutic choice in instances where standard-of-care includes multiple options. A critical step is their application in clinical studies designed to evaluate the capacity to improve selection of drugs for the individual patient. However, the prospective use of microarray technologies and multi-gene models to a single patient sample in the context of a clinical trial requires several methodological advancements.

Methods: We describe a strategy that makes use of a trial-specific reference tumor dataset to provide a source of information for normalization of microarray data between the pre-clinical training set (cell lines) and the single patient sample. Validations of chemosensitivity signatures are repeated under this strategy using historical datasets with split-sample and bootstrap approaches.

Results: Genomic signatures for adriamycin, docetaxel and cisplatin are mapped to Affymetrix GeneChip© HG-U133Plus2.0 platform. Prospective predictions of adriamycin sensitivity were associated with response to neoadjuvant treatment of 133 breast cancers (AUC=0.76, p < 0.001). Prospective predictions of docetaxel (n=14, AUC = 0.92, p = 0.007) and cisplatin sensitivity (n=49, AUC = 0.97, p = 0.0001) discriminated between responders and non-responders under first-line treatment of advanced ovarian cancer. Bootstrap analyses show the genomic signatures require that reference sets have similar characteristics to the investigations sample and be of sufficient size to capture variation in gene expression in the patient population.

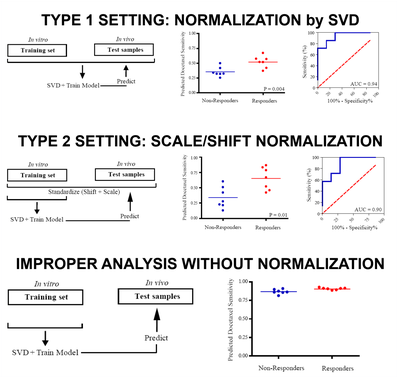
Conclusions: These results demonstrate the accuracy of the genomic signatures are comparable to retrospective assessments when properly applied to a single sample in the prospective setting. We further describe trial designs to test the efficacy of these signatures and an automated data analysis infrastructure to enable their use in a clinical context.

Figure 1: Gene Signatures



Predictors are derived from sensitive and resistant cell lines defined in from Gyorrfy et al and Potti A et al. Genes were selected to be most highly correlated (Pearson) to the cell-line phenotypes. Expression values of each signature are displayed as heatmaps images.

Figure 2: Strategies to Predict Response to Chemotherapy with Cancer Cell-line Data (e.g. Docetaxel)



Evaluating the Gene Signatures in Prospective Clinical Studies

A critical aspect of evaluating the performance of the gene signatures is to define the conditions and procedures necessary for implementation in day-to-day clinical practice. The need for data normalization, as detailed in Figure 2, presents a significant challenge, considering the fact that in the context of a trial, one patient sample must be analyzed at a time.

We propose two different strategies illustrated in Figure 3 that involve three sources of information- 1) the training set data that derives from the cancer cell lines identified as sensitive and resistant to the drug, 2) a reference dataset that provides the tumor specific gene expression context, and 3) the test sample.

As one approach (left panel of Figure 3; Type 1 setting), the training and reference datasets are merged before performing a SVD. The principal components from the merged data set are applied to the Bayesian model without including phenotype information for the reference samples, and the Bayesian model is then applied to post-processed expression data from the test sample.

In an alternative approach (right panel of Figure 3; Type 2 setting), the reference set is used to directly standardize the expression data from each incoming tumor sample to the training set, while determining the principal components and developing the model solely within the training set data.

Figure 3: Prospective Approaches to Gene Signatures for Clinical Applications

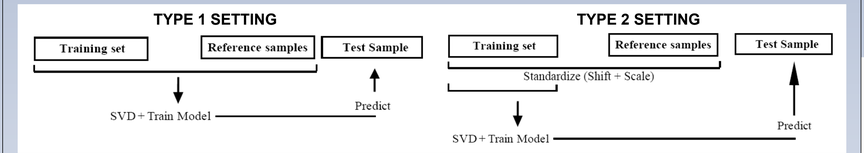
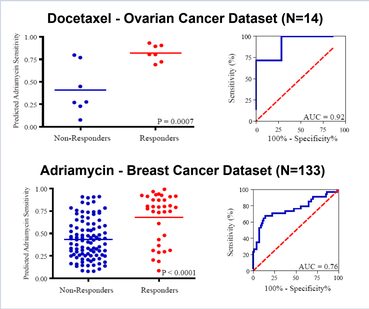
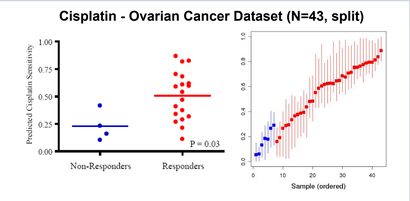


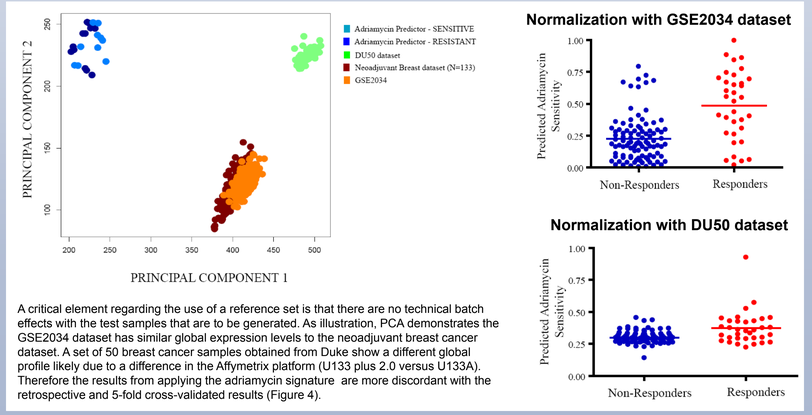
Figure 4: Validation of the Prospective Approaches for Gene Signatures





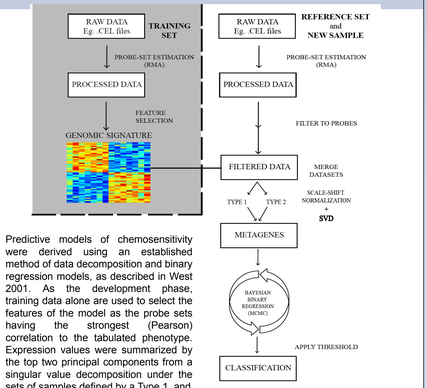
The performance of internal reference sets was evaluated using (A) N = 110 samples from patients that did not receive the investigational agent – Docetaxel; (B) 5-fold cross validation into reference (N = 106/107) and a test (N = 27/26) sets – Adriamycin; and (C) repeated splitting of samples into reference (N=20) and test (N=23) cohorts – Cisplatin. For Cisplatin, the left panel illustrates one iteration; right panel gives the range of results across 100 resamples per subject (average Sens = 0.87, Spec = 0.91). .

Figure 5: Impact of a Reference Set on Predicting Response (e.g. Adriamycin)



A critical element regarding the use of a reference set is that there are no technical batch effects with the test samples that are to be generated. As illustration, PCA demonstrates the GSE2034 dataset has similar global expression levels to the neoadjuvant breast cancer dataset. A set of 50 breast cancer samples obtained from Duke show a different global profile likely due to a difference in the Affymetrix platform (U133 plus 2.0 versus U133A). Therefore the results from applying the adriamycin signature are more discordant with the retrospective and 5-fold cross-validated results (Figure 4).

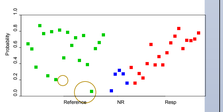
Figure 6: Detailed Work-Flow for Prospective Use of Gene Signatures



Predictive models of chemosensitivity were derived using an established method of data decomposition and binary regression models, as described in West 2001. As the development phase, training data alone are used to select the features of the model as the probe sets having the strongest (Pearson) correlation to the tabulated phenotype. Expression values were summarized by the top two principal components from a singular value decomposition under the sets of samples defined by a Type 1 and Type 2 analysis (see Figures 2 and 3). Summarized expression values (termed 'metagenes' in previous publications) are applied to a Bayesian probit regression model with non-informative priors for the parameters pertaining to the linear model and variance terms. A Markov Chain Monte Carlo (MCMC) is used to obtain the posterior distribution for the linear predictor and regularized probabilities from the training dataset, using 1000 burn-in's and 5000 iterations to ensure convergence. For the validation phase investigational samples are normalized by the SVD or scale/shift standardization, the expected predicted probability is taken as the average value from the posterior distribution of the model.

Figure 7: A Quantitative Approache to Evaluating a Reference Set

Step 2: Apply the prospective algorithm in Fig 6 to each test sample



Step 3: Compute either

(a) deviation from the true phenotype – solid lines, or

(b) misclassification rate for a given threshold – circled points.

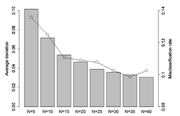
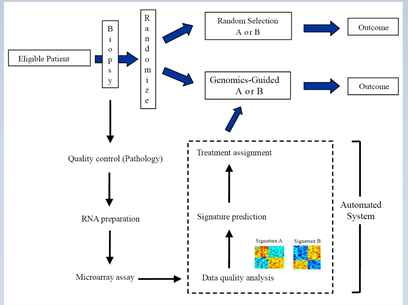


Figure 8: Schema for a Randomized Clinical Trial Design to Evaluate the Performance of Multiple Genomic Signatures in Directing Treatment



Statistical analysis plans for this trial design can provide preliminary evidence of the performance of genomic-guided therapy (aka randomized phase II trial), or definitive evidence of superiority to random selection (aka phase III trial).

References

1. Gyorffy B. et al: Gene expression profiling of 30 cancer cell lines predicts resistance towards 11 anticancer drugs at clinically achieved concentrations. Int. J. Cancer 118: 1699-1712, 2006
2. Potti A. et al: Genomic signatures to guide the use of chemotherapeutics. Nat Med. 12:1294-1300, 2006
3. Hsu DS. et al: Pharmacogenomic strategies provide a rational approach to the treatment of cisplatin-resistant patients with advanced cancer. J. Clin. Oncol. 25:4350-4357. 2007
4. Hess et al: Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin and cyclophosphamide in breast cancer. J. Clin. Oncol. 24:4236-44, 2006
5. West M et al: Predicting the clinical status of human breast cancer by using gene expression profiles. PNAS 98: 11462-7, 2001
6. Bild AH et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 439:353-7, 2006
7. Barry WT et al: Intra-tumor heterogeneity and precision of microarray-based predictors of breast cancer biology and clinical outcome. J Clin Oncol. 28(13), 2198-206, 2010

Duke Comprehensive Cancer Center

VOLUME 28 · NUMBER 13 · MAY 1 2010

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Intratumor Heterogeneity and Precision of Microarray-Based Predictors of Breast Cancer Biology and Clinical Outcome

William T. Barry, Dawn N. Kernagis, Holly K. Dressman, Ryan J. Griffis, J'Vonne D. Hunter, John A. Olson, Jeff R. Marks, Geoffrey S. Ginsburg, Paul K. Marcom, Joseph R. Nevins, Joseph Geradts, and Michael B. Datto

From the Departments of Biostatistics and Bioinformatics, Pathology and Molecular Genetics and Microbiology, and the Institute for Genome Science and Policy, Duke University Medical Center, Durham, NC.

Submitted October 23, 2009; accepted January 26, 2010; published online ahead of print at www.jco.org on April 5, 2010.

Supported by Duke Breast Specialized Program of Research Excellence (SPORE) Grant No. P50CA068438, Congressionally Directed Medical Research Programs Department of Defense Clinical Trial Grant No. W81XWH-07-1-0394 (P.K.M.), and a grant from the V Foundation for Cancer Research.

W.T.B. and D.N.K. contributed equally to this work.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Michael B. Datto, MD, PhD, Department of Pathology, Duke University Medical Center, Medical Center Box 3712, Durham, NC 27710; e-mail: Michael.Datto@duke.edu.

© 2010 by American Society of Clinical Oncology

0732-183X/10/2813-2198/$20.00

DOI: 10.1200/JCO.2009.26.7245

ABSTRACT

Purpose
Identifying sources of variation in expression microarray data and the effect of variance in gene expression measurements on complex predictive and diagnostic models is essential when translating microarray-based experimental approaches into clinical assays. The technical reproducibility of microarray platforms is well established. Here, we investigate the additional impact of intratumor heterogeneity, a largely unstudied component of variance, on the performance of several microarray-based assays in breast cancer.

Patients and Methods
Genome-wide expression profiling was performed on 50 core needle biopsies from 18 breast cancer patients using Affymetrix GeneChip Human Genome U133 Plus 2.0 arrays. Global profiles of expression were characterized using unsupervised clustering methods and variance components models. Array-based measures of estrogen receptor (ER) and progesterone receptor (PR) status were compared with immunohistochemistry. The precision of genomic predictors of ER pathway status, recurrence risk, and sensitivity to chemotherapeutics was evaluated by interclass correlation.

Results
Global patterns of gene expression demonstrated that intratumor variation was substantially less than the total variation observed across the patient population. Nevertheless, a fraction of genes exhibited significant intratumor heterogeneity in expression. A high degree of reproducibility was observed in single-gene predictors of ER (intraclass correlation coefficient [ICC] = 0.94) and PR expression (ICC = 0.90), and in a multigene predictor of ER pathway activation (ICC = 0.98) with high concordance with immunohistochemistry. Substantial agreement was also observed for multigene signatures of cancer recurrence (ICC = 0.71) and chemotherapeutic sensitivity (ICC = 0.72 and 0.64).

Conclusion
Intratumor heterogeneity, although present at the level of individual gene expression, does not preclude precise microarray-based predictions of tumor behavior or clinical outcome in breast cancer patients.

J Clin Oncol 28:2198-2206. © 2010 by American Society of Clinical Oncology

INTRODUCTION

Since its inception, microarray technology has provided a powerful tool to the research community because of its ability to simultaneously measure the expression of tens of thousands of genes. In particular, breast cancer research has seen great benefits from this technology, with many studies describing multigene expression patterns associated with diagnostic and prognostic subclasses among otherwise indistinguishable tumors.[1-3] These studies have also established the ability to predict a cancer patient's treatment response based on gene expression patterns.[4-7] With the clear potential of microarray-based assays to guide clinical decisions, translating these assays to the clinical laboratory is imperative.

Clinical translation requires an understanding of factors that influence the precision and accuracy of microarray-based assays. Chief among these factors is the variability of gene expression measurements, which can be divided into technical (intrinsic to the platform) and preanalytic (intrinsic to the sample) sources of variation. The Affymetrix GeneChip Human Genome U133 Plus 2.0 array platform (Affymetrix, Santa Clara, CA), investigated herein, has a high degree of reproducibility and thus little technical variance, as established by several groups including the MicroArray Quality Control (MAQC)

Information downloaded from jco.ascopubs.org and provided by DUKE MEDICAL LIBRARY SERIALS D on June 1, 2010 from 152.16.179.157.
Copyright © 2010 by the American Society of Clinical Oncology. All rights reserved.

project.[8-12] Variance intrinsic to a sample is more difficult to control, particularly for solid tumor specimens where intratumor heterogeneity could result in significant sampling bias. Small sampling, such as needle core biopsies, can yield samples from the same tumor with different histologic and biologic features. The effect of tumor heterogeneity on microarray-based assays has been evaluated in some cancers, although breast cancer is surprisingly understudied in this regard.[13-17]

Here, we investigate the impact of tumor heterogeneity on several microarray-based predictors of biologic behavior and clinical outcome in breast cancer patients. Multiple core biopsies from individual patients were evaluated by routine histology and tested using single-gene measurements and multigene signatures that would be integral to the routine care of the breast cancer patient, including estrogen receptor (ER) status, progesterone receptor (PR) status, risk of cancer recurrence, and chemotherapeutic sensitivity. Precision for each of these predictors was measured and evaluated in the context of performance expectations for clinical assays.

PATIENTS AND METHODS

Tumor Sample Collection and Histologic Analysis

Following patient consent, samples were obtained from breast cancer excisions as part of a Duke University Health System (DUHS) institutional review board–approved tissue banking and research protocol for the Duke University Medical Center (DUMC) Breast Cancer Specialized Program of Research Excellence (SPORE). Immediately after surgical excision, lumpectomy specimens were sampled by 14-gauge needle core biopsy using an imaging device, as previously described.[18] The core biopsies were embedded in Tissue-Tek OCT (Qiagen, Valencia, CA) and frozen in liquid nitrogen. One 5-μm frozen section was prepared from each sample, stained with hematoxylin and eosin, and evaluated by an expert breast pathologist (J.G.). Routine pathologic evaluation of the corresponding clinical specimens by the DUHS laboratories included determination of ER and PR status by immunohistochemistry (IHC). Samples with an Allred score of 0 or 2 were classified as negative. Human epidermal growth factor receptor 2 (HER2) status was determined by IHC and fluorescent in situ hybridization. This study included only HER2-negative patients with at least two frozen core biopsies containing neoplastic cells.

RNA Purification and Microarray Hybridization

Total RNA was extracted using a kit-based method (RNeasy, Qiagen, Valencia, CA). RNA quality was assessed using an Agilent bioanalyzer (Agilent Technologies, Santa Clara, CA). Hybridization targets were prepared from 2 μg of total RNA and hybridized according to standard Affymetrix protocols using Affymetrix GeneChip Human Genome U133 Plus 2.0 arrays. Arrays were scanned on the Affymetrix GeneChip scanner and probe set expression values, percent present, and 3′/5′ probe set ratios for actin and glyceraldehyde-3-phosphate dehydrogenase were calculated using the Affymetrix Microarray Analysis Suite v5.0.

Microarray Preprocessing and Statistical Analysis

Expression estimates for the 50 DUMC tumor biopsies and for the publicly available breast cancer microarray data sets (GEO# GSE3494[19] and GSE1456[20]) were obtained by robust multiarray average[21] and $\log_2$ transformed. The ratio of intratumor variance to total variance among the DUMC biopsy specimens was calculated for all probe sets. The total variance was calculated as the sum of squared differences from the mean expression for all samples. Intratumor variance was calculated as the sum of squared differences from the mean expression within individual tumors.[14] Global patterns of expression were evaluated by principal component analysis (PCA) and hierarchical clustering using average linkage of the Pearson correlation coefficient.

Predictors of ER pathway activation and breast cancer recurrence were generated from published breast cancer microarray data sets (GEO# GSE3494[19] and GSE1456,[20] respectively), using established methodologies.[22,23] Briefly, tests of differential expression were used to select gene sets strongly correlated to phenotype. Expression values were summarized by the top principal components and fitted to a Bayesian probit regression model. Predicted probabilities were generated for these predictors, and for previously identified signatures of sensitivity to chemotherapeutic agents doxorubicin and docetaxel.[23] Binary classifications were made using thresholds defined a priori for each signature.

The precision of all signatures was evaluated using fixed-effects analysis of variance (ANOVA) models and the intraclass correlation coefficient (ICC [1,1]).[24] ICC values ranging from 0 to 1 have been characterized by Landis and Koch[25] as indicating moderate (0.41-0.60), substantial (0.61-0.80), and almost perfect agreement (0.81-1.00). Accuracy of the single-gene ER and PR predictors and the multigene ER pathway predictor are reported with 95% CIs. The influence of clinical and technical covariates on precision and accuracy are assessed using analysis of covariance (ANCOVA) models. All microarray preprocessing and analysis were performed in R/Bioconductor and Matlab (The Mathworks, Natick, MA) with graphics generated using Graphpad Prism (GraphPad Software, La Jolla, CA) and Cluster/Treeview software (Eisen Lab, Berkeley, CA).[26]

RESULTS

Characterization of Morphologic Heterogeneity in Discrete Samplings of Individual Breast Tumors

Fifty samples were obtained from 18 patients as part of a tissue banking and research protocol for the DUMC Breast Cancer SPORE. One patient had four replicate samples, 12 patients had triplicate samples, and five patients had duplicate samples (Table 1). Analysis was performed on frozen sections stained with hematoxylin and eosin for each of these core biopsies. The sample set contained a mixture of ER- and PR-positive and ER- and PR-negative cases (11 ER-positive/PR-positive, two ER-positive/PR-negative, one ER-negative/PR-positive, four ER-negative/PR-negative) using the diagnostic core biopsy IHC as the standard. Invasive carcinoma was present in 49 samples, while one biopsy contained ductal carcinoma in situ only. One set of biopsies (patient B) contained lobular carcinoma; the remaining cases were ductal-type carcinomas. Invasive tumor cellularity varied from 10% to 90%. RNA and microarray quality control metrics are provided in Table 1.

Intertumor Variance Exceeds Intratumor Variance at the Level of Gene Expression

To evaluate the performance of individual features on the Affymetrix array, we calculated the ratio between intratumor variance and total variance for all probe sets. As shown in Figure 1A to 1C, the majority of samples had an intratumor/total variance ratio below 0.5. A distinct inverse relationship exists between expression intensity and the proportion of intratumor variance. Specifically, within the top quartile of probe sets of highly expressed genes (Fig 1A) more than 90% had an intratumor variance ratio less than one third of the total variance. Conversely, only 11% of genes expressed at comparatively low levels (bottom quartile, Fig 1A) had an intratumor/total variance ratio less than one third. We also found that probe sets corresponding to the U133B platform generally had higher intratumor variability when compared with those on the original U133A array ($P < .001$; Fig 1B), and a similar pattern was noted for probe sets annotated to known genes compared with unannotated probe sets ($P < .001$).

© 2010 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by DUKE MEDICAL LIBRARY SERIALS D on June 1, 2010 from 152.16.179.157.
Copyright © 2010 by the American Society of Clinical Oncology. All rights reserved.

Table 1. Histology, RNA, and Microarray Quality Control Measures

	Histology QC								RNA QC			Microarray QC		
Patient Sample	Histology	Size	Grade	ER	PR	HER2	CA (%)	Other Features	Concentration (μg/μL)	260/280	260/230	% P	Actin	GAPDH
A1	Ductal	2.3	3	+	+	−	90		0.339	2.08	2.03	54.2	1.49	1.44
A2							80		0.097	2.09	2.37	49.7	2.09	2.09
A3							65		0.162	2.09	2.08	53.9	2.04	1.77
B1	Lobular	3.6	2	+	+	−	50		0.068	2.02	2.24	53.4	1.34	1.26
B2							85		0.08	2.09	2.44	50.6	1.21	1.12
B3							70		0.068	2.11	1.86	53	1.81	1.36
C1	Ductal	2	2	+	+	−	10	Biopsy site	0.042	1.98	2.39	54.3	1.49	1.14
C2							80	Biopsy site	0.093	2.06	1.63	56	1.37	1.01
C3							40	Biopsy site	0.07	2.04	0.84	57	1.24	1
D1	Ductal	2.8	2	+	+	−	30		0.158	2.08	2.27	55.3	1.54	1.31
D2							50		0.147	2.08	1.59	50.4	1.35	1.28
D3							70		0.182	2.1	2.22	52	1.21	1.23
E1	Ductal	2.1	3	−	−	−	25		0.269	2.08	1.82	52.7	1.17	1.03
E2							75		0.324	2.08	2.27	50.9	1.14	1.05
E3							50		0.26	2.08	2.2	53.5	1.08	1.09
F1	Ductal/micropapillary	Multi-focal	2	+	−	−	25		0.06	2.11	1.8	54.1	1.89	1.17
F2							60		0.292	2.06	2.19	54	1.19	1.13
F3							25		0.079	2.04	2.13	54.2	1.18	0.99
G1	Ductal	3.4	3	−	−	−	70		0.515	2.08	2.19	51.3	3.31	1.31
G2							30		0.076	2.03	2.39	56.1	1.13	0.98
G3							70		0.331	2.07	2.24	54.4	1.21	1.04
H1	Ductal	1.6	3	−	−	−	30		0.112	2.11	1.59	50.3	2.94	1.12
H2							30		0.096	1.92	0.92	54.1	2.21	1.09
H3							30		0.072	2.12	1.29	49.7	1.49	1.09
H4							20		0.285	2.09	1.98	49.6	6.66	1.95
I1	Ductal/lobular	>2	2	+	+	−	70		0.237	2.11	1.96	52.5	2.35	1.19
I2							70		0.331	2.09	2.03	51	2.17	1.15
I3							70		0.173	2.12	2.07	51.5	1.79	1.1
J1	Ductal	3.5	3	+	+	−	80		0.682	2.08	1.93	49.8	3.85	1.47
J2							80		0.941	2.08	2.11	51.3	2.91	1.42
J3							80		0.561	2.07	2.17	50.8	3.49	1.57
K1	Ductal	2.4	2	+	+	−	15	DCIS	0.068	2.12	1.66	51.4	1.88	1.22
K2							75	DCIS	0.044	2.06	1.85	55	1.85	1.04
L1	Ductal	1.3	3	+	+	−	0	DCIS	0.055	2.1	1.31	51.2	1.55	1.17
L2							15	DCIS	0.083	2.11	1.84	50.7	1.39	0.92
L3							15	DCIS	0.099	2.09	1.21	51	1.69	1.06
M1	Ductal	1.9	2	+	+	−	50		0.116	2.1	1.74	50.3	2.49	1.75
M2							50	Necrosis	0.09	2.08	1.74	49.1	1.58	1.1
M3							30		0.084	2.12	1.65	45.5	2.06	1.35
N1	Ductal	1.8	2	+	+	−	20		0.077	2.14	1.8	50.5	1.57	1.19
N2							70		0.068	2.06	1.96	52.9	1.28	1.09
N3							70		0.069	2.13	0.67	51.6	1.74	1.33
O1	Ductal	3	1	+	+	−	20	Carcinoma + papilloma	0.261	2.11	2.2	53.7	2	0.95
O2							50	Carcinoma + papilloma	0.143	2.13	1.24	55.5	2.06	1.13
P1	Ductal	0.9	3	−	−	−	60		0.05	2.11	1.89	54.4	1.34	1.06
P2							10		0.044	2.15	1.61	56.6	1.34	1.05
Q1	Ductal	1.8	3	−	+	−	70	Biopsy site	0.474	2.07	2.03	51.9	2.13	1.08
Q2							30	Biopsy site	0.226	2.11	2.03	55.1	2.02	1.02
R1	Ductal	2.8	3	+	−	−	70	Biopsy site	0.156	2.07	2.1	52.8	1.66	1.12
R2							20	Biopsy site	0.132	2.1	2.14	52.8	2.33	1.26

NOTE. 260/280 and 260/230 are optical density ratios; CA is the tumor content in percent; % P is the percent of probe sets present; actin and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) are the 3′ to 5′ probe set ratios for the actin and GAPDH probe sets, respectively.

Abbreviations: QC, quality control; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; DCIS, ductal carcinoma-in-situ.

To evaluate the global patterns of expression, we performed PCA and hierarchical clustering on robust multiarray average–normalized data. Hierarchical clustering of all samples using probe sets with $\log_2$ expression values > 5 demonstrated that replicate samples from a single tumor tended to group together in a robust and statistically significant fashion (Fig 2A and Data Supplement, online only). One tumor showed imperfect clustering (patient F). There were no histology, RNA, or array quality metrics that could account for discordance

Information downloaded from jco.ascopubs.org and provided by DUKE MEDICAL LIBRARY SERIALS D on June 1, 2010 from 152.16.179.157.
Copyright © 2010 by the American Society of Clinical Oncology. All rights reserved.

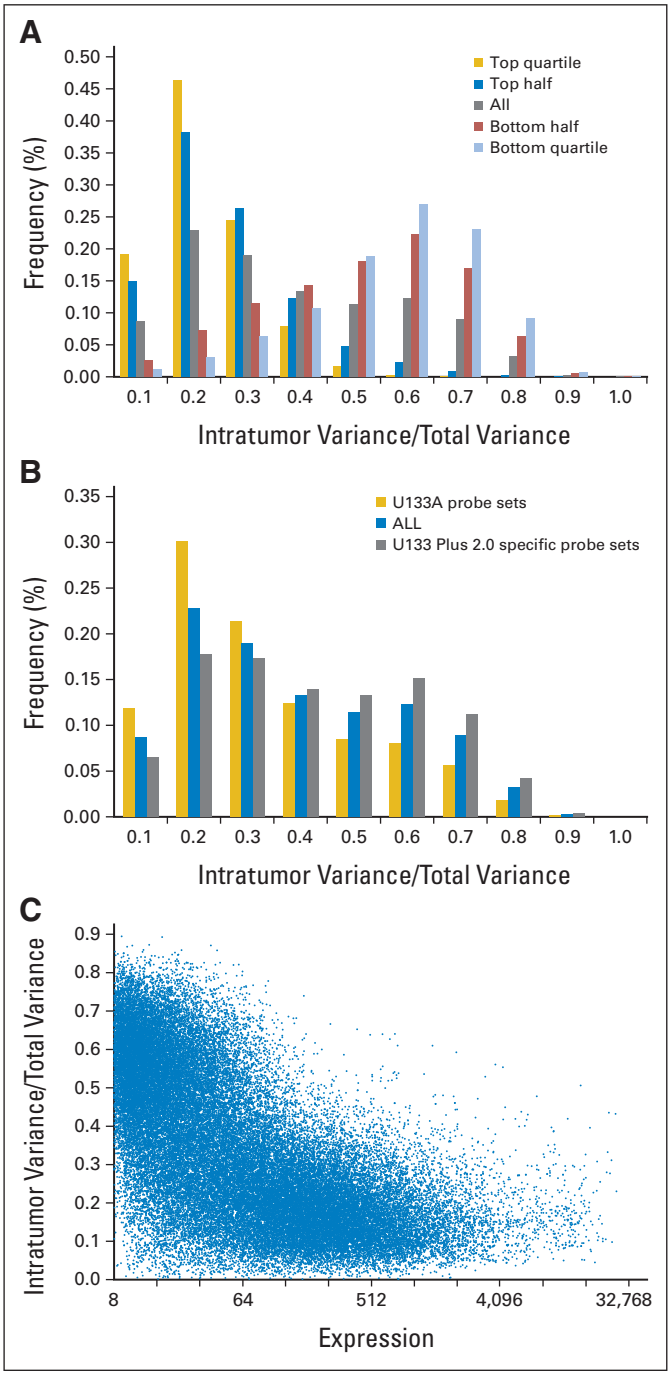


Fig 1. For Affymetrix probe sets, intratumor variance is generally a small component of total variance. Histograms stratified by (A) expression level demonstrate that profiles differ sharply between highly and lowly expressed genes, and (B) source demonstrate that U133A probe sets are generally more reproducible. (C) Scatter plot demonstrates an inverse relationship between variance ratios and average expression.

between samples F1, F2, and F3. A PCA of the filtered expression data (Fig 2B) demonstrates that patient replicates largely cluster together in the top two principal components (capturing 28% of the total variance in global expression), indicating that more heterogeneity is seen across patients than within replicates. In summary, while most genes exhibit a low degree of intratumor variability and most replicate samples demonstrate similar global expression patterns, poor performing probe sets exist and could potentially have an impact on the precision of array-based predictors of tumor biology or behavior.

Intratumor Variance in Gene Expression Does Not Preclude Precise Predictions of Tumor Biology

ER and PR status are two critical characteristics of breast cancer that define biologically distinct subgroups of disease. ER status and PR status are prognostic of clinical outcome and often determine the course of treatment. To determine the effect of intratumor variance on array-based assessments of hormone receptor status, we began by evaluating single-gene predictors of ER (probe set 205225_at)[27] and PR expression (probe set 208305_at). Expression levels of 205225_at demonstrated almost perfect agreement among the replicates samples (ICC, 0.94; 95% CI, 0.86 to 0.97; Fig 3A). Further, using an optimal binary classification (threshold of $\log_2 = 9.6$ to maximize accuracy), only one patient showed discordance with IHC (sensitivity, 1.0; 95% CI, 0.90 to 1.0; specificity, 0.88; 95% CI, 0.62 to 0.98). A single probe set for the PR gene, 208305_at, also showed near perfect agreement among replicates (ICC, 0.95; 95% CI, 0.89 to 0.98), but with an optimal binary classifier (threshold of $\log_2 = 5.2$ chosen to maximize accuracy), greater disagreement was noted between PR expression and IHC results. While all samples from the six PR-negative patients were classified correctly (specificity, 1.0; 95% CI, 0.80 to 1.0), samples from five of the 12 PR-positive patients were discordant with IHC (sensitivity, 0.64; 95% CI, 0.45 to 0.80).

We next created a multigene predictor of ER pathway activation from a previously published breast cancer Affymetrix U133A microarray data set (GEO# GSE3494). This data set was filtered to retain probe sets with mean $\log_2$ expression values > 5.0 for the 247 ER-annotated patient samples.[19] The ER predictor was based on 1,022 probe sets identified by a Wilcoxon rank sum test with Bonferroni correction (adjusted $P < .05$), with intentional exclusion of probe sets for ER itself (Data Supplement). The large number of differentially expressed genes highlights the distinct biologic characteristics of these different tumor types. A predictor of ER pathway status was generated by applying the top principal component of expression from the 1,022 probe set list to a Bayesian probit regression model. Under leave-one-out cross validation and a threshold of 0.5, the model classified 91% of the ER-negative samples and 85% of the ER-positive samples by IHC correctly (Fig 3B). Applied prospectively to the breast cancer replicate data set with an optimal threshold of 0.45 (Fig 3C), the model showed near perfect precision (ICC, 0.98; 95% CI, 0.95 to 0.99) and 96% accuracy identical to the single-gene model. Taken together, these data show that when assaying a robust biologic property of breast cancer such as ER status, intratumor variance does not preclude precise predictions from microarray data.

Intratumor Heterogeneity Does Not Preclude Precise Predictions of Clinical Outcome

A prognostic model for death attributed to breast cancer was generated from expression data from a previously published data set (GEO# GSE1456; N = 159).[20] Probe sets associated with survival were identified using a Cox proportional hazard model and the Benjamini-Hochberg method for controlling false discovery rates.[28] A total of 205 probe sets (representing 184 genes) were identified with a false discovery rate < 0.01. These probe sets contain an over-representation of genes involved in cell cycles (20 genes; $P < .001$), cytokinesis (seven genes; $P = .004$), and cellular metabolism (84 genes; $P = .004$; Data

© 2010 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by DUKE MEDICAL LIBRARY SERIALS D on June 1, 2010 from 152.16.179.157.
Copyright © 2010 by the American Society of Clinical Oncology. All rights reserved.

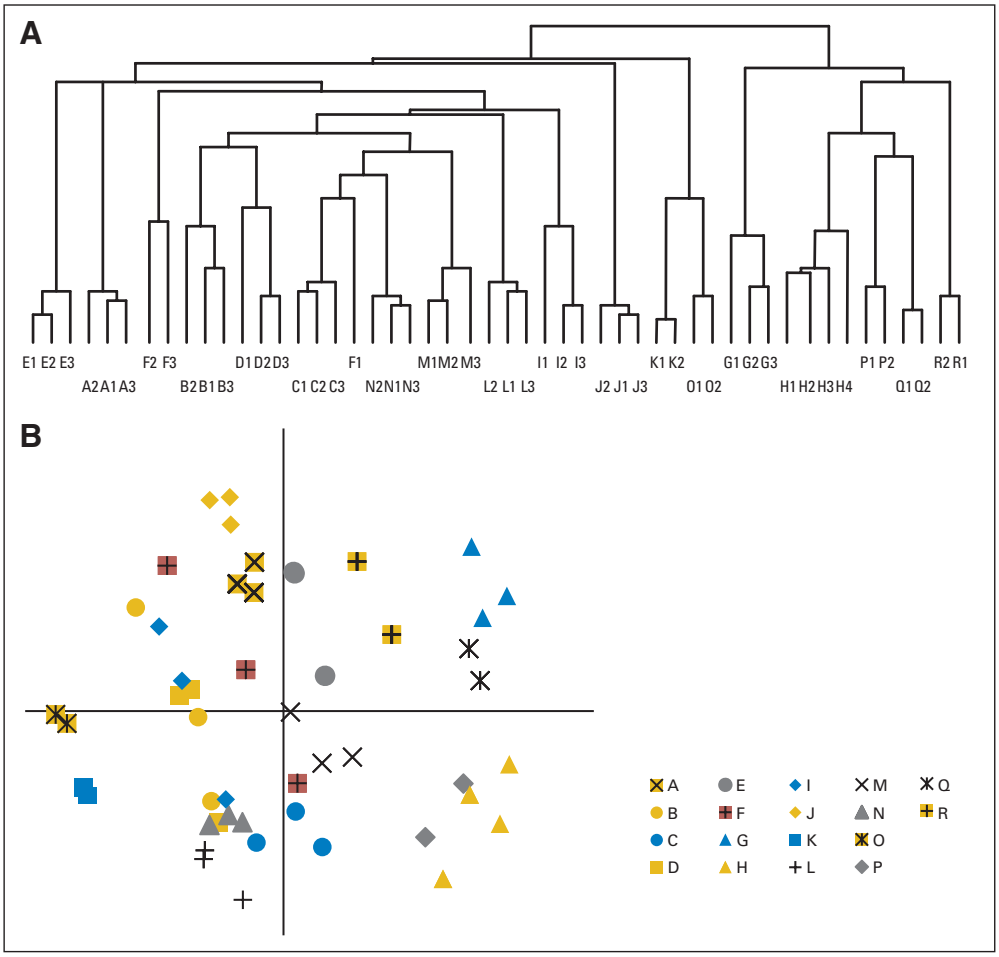


Fig 2. Unsupervised clustering demonstrates that replicate samples have similar global patterns of expression. (A) Pearson Correlation Coefficient–based hierarchical clustering shows complete segregation of replicates in 17 of 18 patients and is confirmed by (B) a scatter plot of the top two principal components of all expressed genes. The discordant patient, F, is highlighted in red. Principal component 1 captures estrogen receptor status.

Supplement). A binary classifier of survival was created using the first two principal components. The fitted model accurately stratified patients into high and low risk for death from disease with a threshold of 0.50 (Fig 4A). When this genomic signature was next applied to an independent validation data set (GSE3494; N = 315),[19] it maintained the ability to identify a high-risk cohort for survival ($P = .0069$; Fig 4B). The precision of the prognostic genomic signature was found to be substantial (ICC, 0.71; 95% CI, 0.48 to 0.87), when applied to the replicate samples (Fig 4C). The patient with the most variance in recurrence risk predictions (patient F) also showed the poorest internal concordance by hierarchical clustering and PCA analysis, suggesting that the global variance in expression within this tumor may have affected the precision of the risk predictor.

Finally, we established the precision of previously published chemotherapeutic sensitivity predictors for doxorubicin and docetaxel (Fig 5A and Data Supplement), as implemented in a randomized phase II trial to direct neoadjuvant therapy.[23] A substantial level of agreement is observed for the doxorubicin sensitivity predictor (ICC, 0.72; 95% CI, 0.49 to 0.88). Further, 12 of 18 patients showed complete concordance when applying an a priori threshold of 0.6 used in the clinical trial. The docetaxel sensitivity predictions showed a slightly lower level of precision (ICC, 0.64; 95% CI, 0.38 to 0.83), with 14 of 18 patients showing complete concordance. Under a multilevel classification of higher sensitivity to one agent or of double resistance (Fig 5B), 13 of 18 patients showed complete concordance. By resampling from a binomial mixture distribution, the observed agreement in the three-level classification was highly significant ($P < .001$). Variation in tumor content and quality control measures for RNA and Affymetrix arrays was found not to be associated with the discordance in the multigene predictors of clinical outcome (all ANCOVA $P > .05$).

DISCUSSION

Here, we demonstrate that intratumor variance at the level of gene expression does not preclude the development of precise microarray-based clinical prediction models in breast cancer. We show intertumor heterogeneity is greater than intratumor heterogeneity at the level of global gene expression for this breast tumor data set. While a small group of genes exhibits a significant level of intratumor variation, many of these genes are expressed at relatively low levels and can be filtered as background noise when creating predictive algorithms. Finally, we demonstrate that a high degree of precision was seen among replicate samples when assayed using single-gene predictors of ER and PR status and PCA-based predictors of ER pathway activation, cancer recurrence, and chemotherapeutic sensitivity.

The heterogeneous nature of solid neoplasms has been recognized and studied by pathologists for decades. In the context of breast

 © 2010 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by DUKE MEDICAL LIBRARY SERIALS D on June 1, 2010 from 152.16.179.157.
Copyright © 2010 by the American Society of Clinical Oncology. All rights reserved.

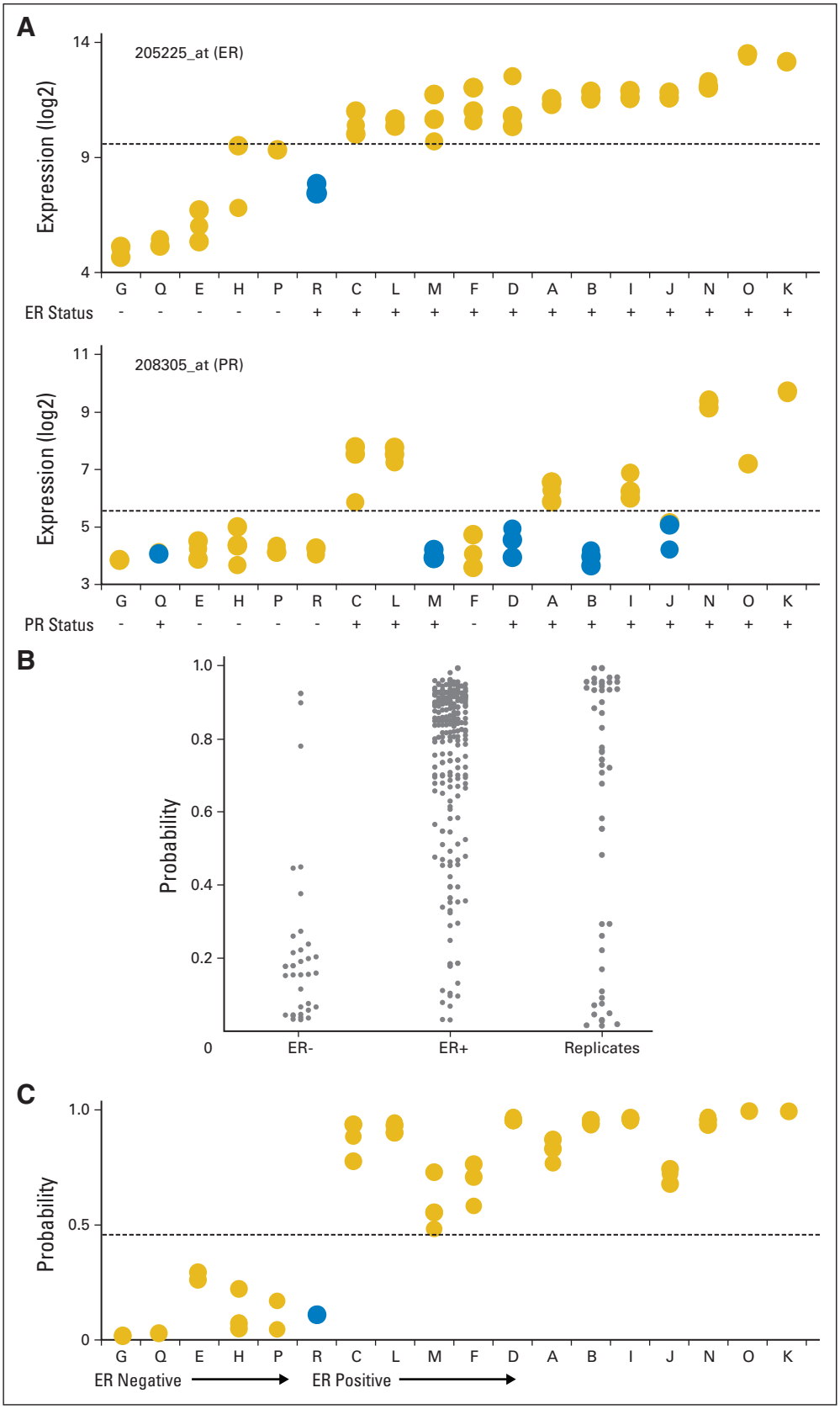


Fig 3. (A) Single-gene predictors of estrogen receptor (ER) and progesterone receptor (PR) status demonstrate high precision among individual patients and agreement with immunohistochemistry. (B) A multigene predictor of ER pathway activation generated from an independent data set of 247 patients (C) demonstrates high precision and agreement with immunohistochemistry in the breast replicate data set. Blue solid circles, discordant or inaccurate samples; gold solid circles, complete concordance. Capital letters are patient designations.

© 2010 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by DUKE MEDICAL LIBRARY SERIALS D on June 1, 2010
from 152.16.179.157.
Copyright © 2010 by the American Society of Clinical Oncology. All rights reserved.

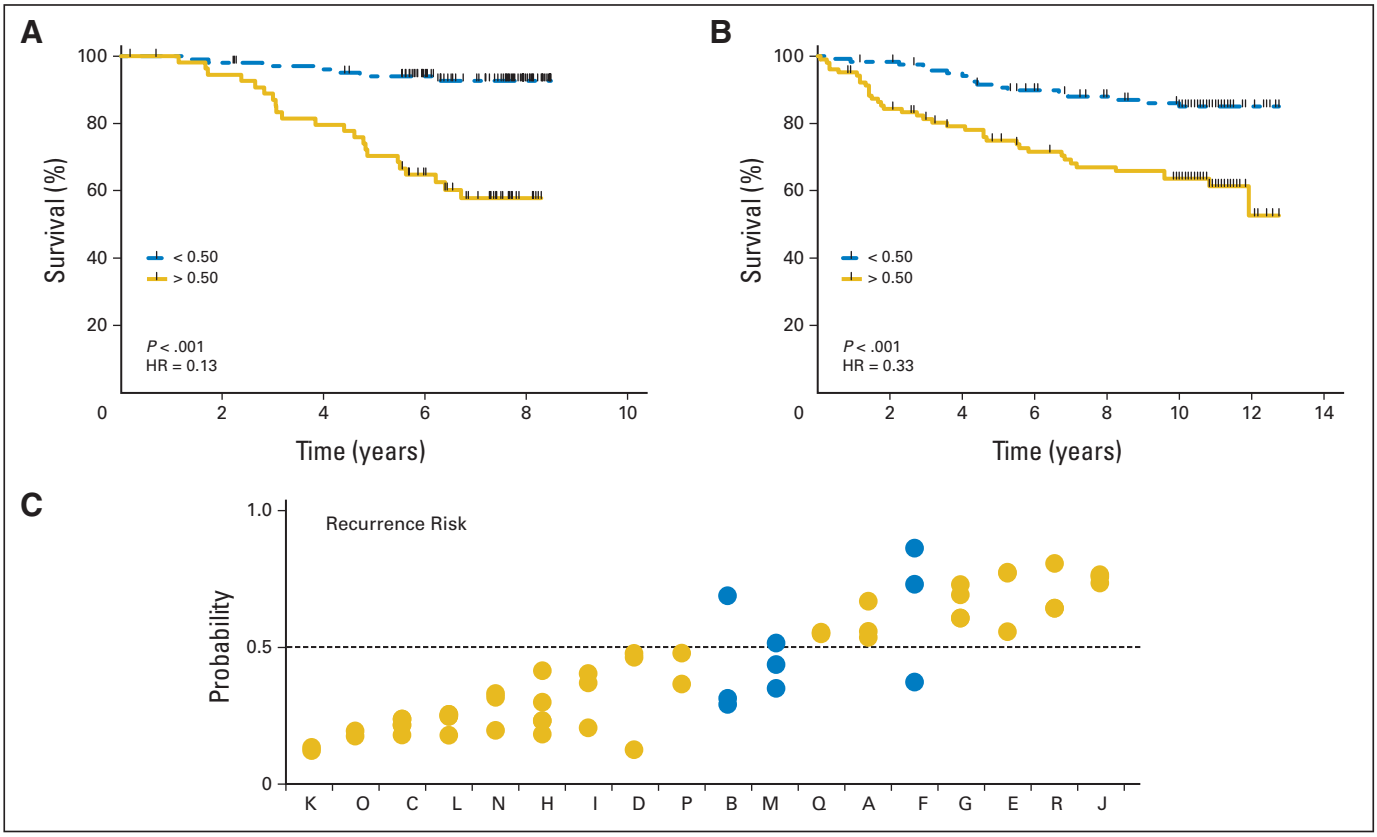


Fig 4. Novel multigene signatures for breast cancer survival demonstrate prognostic value by Kaplan-Meier plot and log-rank tests under (A) cross-validation in the training data (GSE1456; N = 159), (B) independent validation set in a second data set (GSE3494; N = 315), and (C) substantial agreement when applied to the replicate data set. Blue solid circles, discordant or inaccurate samples; gold solid circles, complete concordance. Capital letters are patient designations. HR, hazard ratio.

cancer, the varying presence of normal breast tissue, inflammatory cells, vessels, necrosis, and neoplastic epithelium gives rise to a variably mixed population of cells with unique or distinct biologic makeup in any given tumor sampling. Thus, the question has been raised of how this cellular heterogeneity may affect assays typically performed for diagnostic and prognostic purposes in breast cancer patients, including ER, PR, and HER2.

The largest study of the reproducibility of HER2 testing was performed on patients enrolled in the North Central Cancer Treatment Group (NCCTG) N9831 adjuvant trial of trastuzumab. In this study, only 85.8% of the 2,535 patients registered in the trial had concordant results for HER2 positivity between the local and central performing laboratories (88.1% concordance by fluorescent in situ hybridization and 81.6% concordance by IHC).[29] A similar study of ER and PR IHC measurements in 776 patients enrolled on Eastern Cooperative Oncology Group (ECOG) 2197 showed 90% and 84% concordance between local and central laboratory studies for ER and PR status, respectively.[30] These studies, however, are strictly a measure of assay reproducibility; a measure of technical variance rather than tumor heterogeneity.

Discordance attributable solely to tumor heterogeneity between breast core biopsy and resection specimens ranges from 1.2% for ER status on the low end[31] to 14% for ER status and 20% for PR status on the high end.[32] Discordance rates attributable to technique and tumor heterogeneity that develops over time was determined to be even higher (18.4%, 40.3%, and 13.6% for ER, PR, and HER2, respectively) when comparing primary tumors and their corresponding metastatic lesions.[33] The low discordance rates (0% in this study) of our array-based predictors of ER and PR status are at least equivalent to the good performance of these traditional techniques.

While the effect of tumor heterogeneity and technical variance on ER, PR, and HER2 testing has been well studied, few studies have examined the effect of tumor heterogeneity on multigene predictive algorithms. A study comparable to the one presented here reached similar conclusions using a 48-gene TaqMan-based assay.[34] This study showed high concordance among three replicate samplings for 12 breast cancer patients. A similar study examining the contribution of technical variance to the reproducibility of microarray-based assays in breast cancer, demonstrated that gene expression–based signatures developed from replicate experiments among 35 patients resulted in precise predictions of breast cancer chemotherapeutic responsiveness.[13]

To the best of our knowledge, our work is the first to demonstrate the impact of intratumor heterogeneity on the precision of multigene predictive models focused on tumor behavior. Precision is an integral component of clinical testing that is often overlooked in the early stages of translational research. Accuracy always seems of primary interest at that stage. However, our data suggest that a lack of accuracy in microarray-based assays may in fact be caused by a lack of precision, particularly for more indirect measures of tumor behavior (eg, chemosensitivity or recurrence). For these more abstract measures, a lack

 © 2010 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by DUKE MEDICAL LIBRARY SERIALS D on June 1, 2010
from 152.16.179.157.
Copyright © 2010 by the American Society of Clinical Oncology. All rights reserved.

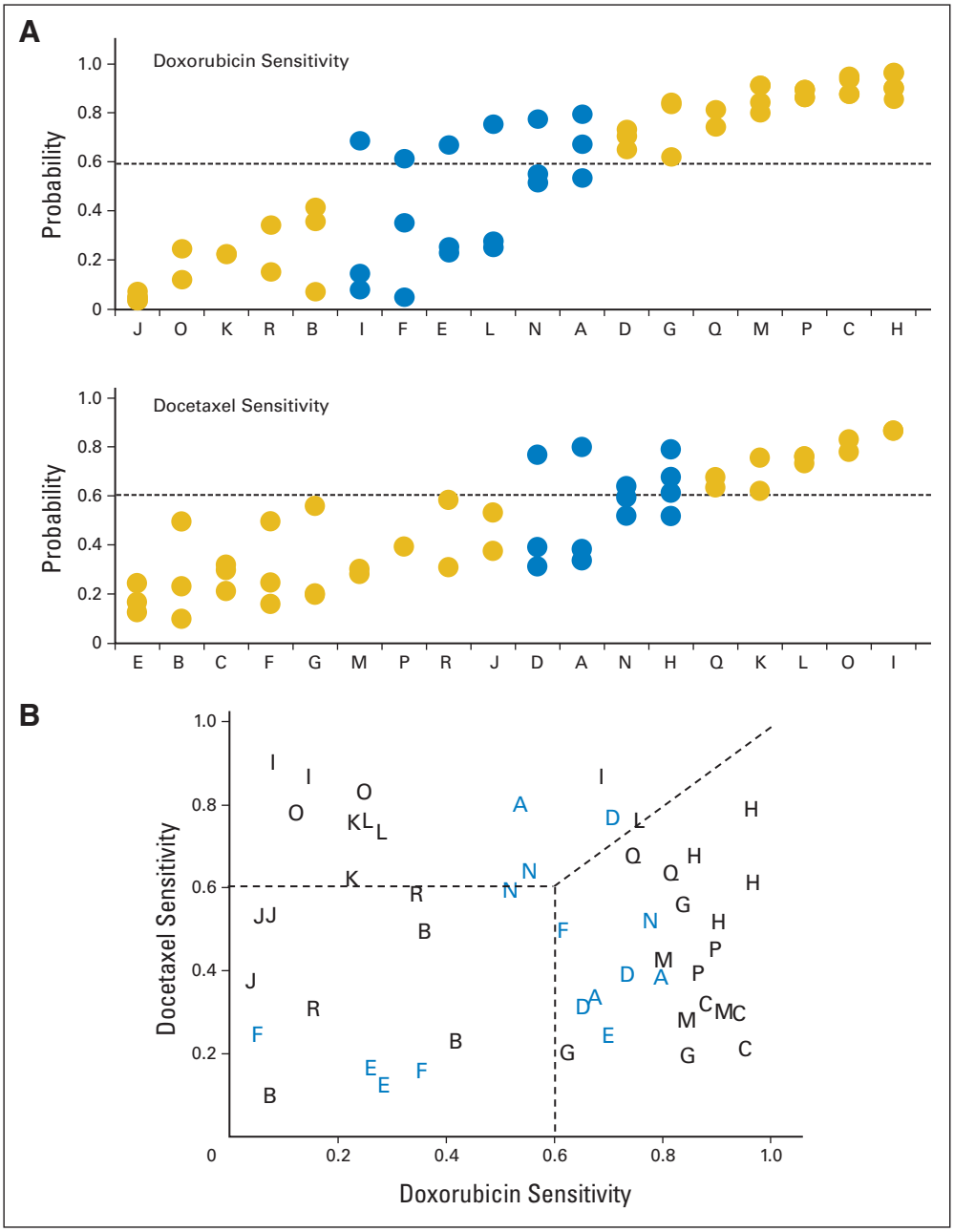


Fig 5. (A) Multigene signatures of chemosensitivity to doxorubicin and docetaxel demonstrate substantial agreement when applied prospectively to the replicate samples (A-R) breast cancer data set. Capital letters are patient designations. Gold solid circles, complete concordance; blue solid circles, discordant samples. (B) Under a multilevel classification of higher sensitivity to one agent or of double resistance, 13 of 18 patients show complete concordance.

of precision may be attributable to true differences in tumor microenvironment. This is supported by our finding that the genes in these complex predictors tend to show higher intratumor variance (Data Supplement). In fact, there is a direct correlation between each predictor's ICC and the proportion of genes with high intratumor variance. Our data also highlight the fact that testing of replicate samples, even early in the development of a microarray-based assay, can clearly differentiate between inaccurate and imprecise.

The importance of assessing performance of multigene microarray-based assays, as described here, bears on the future use of this technology as a single assay for breast cancer patients that can provide not only measures of prognosis or predicted therapeutic response, but can also supplement or replace the standard assays of ER, PR, and HER2. The feasibility of this approach is, in part, demonstrated in this study. This full genome expression profile may find additional uses as assays become algorithms applied to expression data sets. As we move toward this, understanding the role of tumor heterogeneity in measures of tumor behavior and developing approaches and data sets (like the one provided here) to test the precision of these algorithms is essential.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships

© 2010 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by DUKE MEDICAL LIBRARY SERIALS D on June 1, 2010 from 152.16.179.157.
Copyright © 2010 by the American Society of Clinical Oncology. All rights reserved.

marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.
Employment or Leadership Position: John A. Olson, Core Prognostex (U) **Consultant or Advisory Role:** Joseph R. Nevins, Millennium Pharmaceuticals (C), Johnson & Johnson (C) **Stock Ownership:** John A. Olson, Core Prognostex; Joseph R. Nevins, Expression Analysis **Honoraria:** Joseph R. Nevins, Eli Lilly **Research Funding:** None **Expert Testimony:** None **Other Remuneration:** Michael B. Datto, Affymetrix

AUTHOR CONTRIBUTIONS

Conception and design: William T. Barry, John A. Olson, Jeff R. Marks, Geoffrey S. Ginsburg, Paul K. Marcom, Joseph R. Nevins, Joseph Geradts, Michael B. Datto
Financial support: Paul K. Marcom, Joseph R. Nevins
Administrative support: Geoffrey S. Ginsburg, Joseph R. Nevins
Provision of study materials or patients: Paul K. Marcom, Joseph Geradts
Collection and assembly of data: William T. Barry, Holly K. Dressman, Ryan J. Griffis, J'Vonne D. Hunter, Joseph Geradts, Michael B. Datto
Data analysis and interpretation: William T. Barry, Dawn N. Kernagis, Jeff R. Marks, Paul K. Marcom, Joseph R. Nevins, Joseph Geradts, Michael B. Datto
Manuscript writing: William T. Barry, Dawn N. Kernagis, Paul K. Marcom, Joseph R. Nevins, Joseph Geradts, Michael B. Datto
Final approval of manuscript: William T. Barry, Dawn N. Kernagis, Holly K. Dressman, Ryan J. Griffis, J'Vonne D. Hunter, John A. Olson, Jeff R. Marks, Geoffrey S. Ginsburg, Paul K. Marcom, Joseph R. Nevins, Joseph Geradts, Michael B. Datto

REFERENCES

1. Perou CM, Sorlie T, Eisen MB, et al: Molecular portraits of human breast tumours. Nature 406: 747-752, 2000

2. Rouzier R, Perou CM, Symmans WF, et al: Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 11:5678-5685, 2005

3. van de Vijver MJ, He YD, van't Veer LJ, et al: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347:1999-2009, 2002

4. Potti A, Nevins JR: Utilization of genomic signatures to direct use of primary chemotherapy. Curr Opin Genet Dev 18:62-67, 2008

5. Sotiriou C, Pusztai L: Gene-expression signatures in breast cancer. N Engl J Med 360:790-800, 2009

6. Bonnefoi H, Potti A, Delorenzi M, et al: Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: A substudy of the EORTC 10994/BIG 00-01 clinical trial. Lancet Oncol 8:1071-1078, 2007

7. Salter KH, Acharya CR, Walters KS, et al: An integrated approach to the prediction of chemotherapeutic response in patients with breast cancer. PLoS ONE 3:e1908, 2008

8. Canales RD, Luo Y, Willey JC, et al: Evaluation of DNA microarray results with quantitative gene expression platforms. Nat Biotechnol 24:1115-1122, 2006

9. Patterson TA, Lobenhofer EK, Fulmer-Smentek SB, et al: Performance comparison of one-color and two-color platforms within the MicroArray Quality Control (MAQC) project. Nat Biotechnol 24:1140-1150, 2006

10. Shi L, Reid LH, Jones WD, et al: The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 24:1151-1161, 2006

11. Irizarry RA, Warren D, Spencer F, et al: Multiple-laboratory comparison of microarray platforms. Nat Methods 2:345-350, 2005

12. Zakharkin SO, Kim K, Mehta T, et al: Sources of variation in Affymetrix microarray experiments. BMC Bioinformatics 6:214, 2005

13. Anderson K, Hess KR, Kapoor M, et al: Reproducibility of gene expression signature-based predictions in replicate experiments. Clin Cancer Res 12:1721-1727, 2006

14. Bachtiary B, Boutros PC, Pintilie M, et al: Gene expression profiling in cervical cancer: An exploration of intratumor heterogeneity. Clin Cancer Res 12:5632-5640, 2006

15. Francis P, Fernebro J, Eden P, et al: Intratumor versus intertumor heterogeneity in gene expression profiles of soft-tissue sarcomas. Genes Chromosomes Cancer 43:302-308, 2005

16. Jochumsen KM, Tan Q, Hølund B, et al: Gene expression in epithelial ovarian cancer: A study of intratumor heterogeneity. Int J Gynecol Cancer 17: 979-985, 2007

17. Trautmann K, Steudel C, Grossmann D, et al: Expression profiling of gastric cancer samples by oligonucleotide microarray analysis reveals low degree of intra-tumor variability. World J Gastroenterol 11:5993-5996, 2005

18. Olson JJ, Morris E, Zee KV, et al: Magnetic resonance imaging facilitates breast conservation for occult breast cancer. Ann Surg Oncol 7:411-415, 2000

19. Miller LD, Smeds J, George J, et al: An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. Proc Natl Acad Sci U S A 102:13550-13555, 2005

20. Pawitan Y, Bjöhle J, Amler L, et al: Gene expression profiling spares early breast cancer patients from adjuvant therapy: Derived and validated in two population-based cohorts. Breast Cancer Res 7:R953–R964, 2005

21. Irizarry R, Hobbs B, Collin F, et al: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4:249-264, 2003

22. West M, Blanchette C, Dressman H, et al: Predicting the clinical status of human breast cancer by using gene expression profiles. Proc Natl Acad Sci U S A 98:11462-11467, 2001

23. Potti A, Dressman H, Bild A, et al: Genomic signatures to guide the use of chemotherapeutics. Nat Med 12:1294-1300, 2006

24. Shrout P, Fleiss J: Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 86: 420-428, 1979

25. Landis J, Koch G: The measurement of observer agreement for categorical data. Biometrics 33:159-174, 1977

26. Eisen MB, Spellman PT, Brown PO, et al: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863-14868, 1998

27. Gong Y, Yan K, Lin F, et al: Determination of oestrogen-receptor status and ERBB2 status of breast carcinoma: A gene-expression profiling study. Lancet Oncol 8:203-211, 2007

28. Benjamini Y, Drai D, Elmer G, et al: Controlling the false discovery rate in behavior genetics research. Behav Brain Res 125:279-284, 2001

29. Perez EA, Suman VJ, Davidson NE, et al: HER2 testing by local, central, and reference laboratories in specimens from the North Central Cancer Treatment Group N9831 intergroup adjuvant trial. J Clin Oncol 24:3032-3038, 2006

30. Badve SS, Baehner FL, Gray RP, et al: Estrogen- and progesterone-receptor status in ECOG 2197: Comparison of immunohistochemistry by local and central laboratories and quantitative reverse transcription polymerase chain reaction by central laboratory. J Clin Oncol 26:2473-2481, 2008

31. Hodi Z, Chakrabarti J, Lee AH, et al: The reliability of assessment of oestrogen receptor expression on needle core biopsy specimens of invasive carcinomas of the breast. J Clin Pathol 60:299-302, 2007

32. Taucher S, Rudas M, Gnant M, et al: Sequential steroid hormone receptor measurements in primary breast cancer with and without intervening primary chemotherapy. Endocr Relat Cancer 10:91-98, 2003

33. Liedtke C, Broglio K, Moulder S, et al: Prognostic impact of discordance between triple-receptor measurements in primary and recurrent breast cancer. Ann Oncol 20:1953-1958, 2009

34. Lyng M, Laenkholm A, Pallisgaard N, et al: Intratumor genetic heterogeneity of breast carcinomas as determined by fine needle aspiration and TaqMan low density array. Cell Oncol 29:361-372, 2007

Information downloaded from jco.ascopubs.org and provided by DUKE MEDICAL LIBRARY SERIALS D on June 1, 2010 from 152.16.179.157.
Copyright © 2010 by the American Society of Clinical Oncology. All rights reserved.

Pathologic response analysis for a genomically guided preoperative chemotherapy trial

P. K. Marcom[4,5], J. Olson Jr.[3,5], J. R. Nevins[5], V. Liotcheva[5], T. Foster[5], G. Ginsburg[5], J. Marks[3,5], M. Datto[2,5], W. T. Barry[1,5], J. Geradts[2]

Duke Multidisciplinary Breast Program

1)Dept. of Biostatistics and Bioinformatics, 2) Dept. of Pathology, 3) Dept. of Surgery, 4) Div. of Oncology , 5) Institute for Genome Sciences and Policy
Duke University

Email: marco001@mc.duke.edu

ABSTRACT

Background: "Performance of Genomic Expression Profiles to Direct the Use of Preoperative Chemotherapy for Early Stage Breast Cancer" was a prospective randomized trial designed to assess in vitro derived signatures for predicting sensitivity to doxorubicin and docetaxel. Methods: Patients with HER2 negative early-stage breast greater than 1.5 cm were eligible. Either banked or fresh cores were then profiled using the Affymetrix U133Plus2.0 microarray. Patients with adequate microarray data were then randomized in a 2:1 ratio to a genomically-guided arm using the microarray data to assign treatment with either doxorubicin (AC) or docetaxel (TC) based chemotherapy or a control arm using random assignment to AC or TC (standard doses every 3 weeks for 4 cycles). Pathologic complete response (pCR) was defined as no residual invasive cancer. The primary aims were to determine whether genomic profiling could improve pCR rates compared to random assignment; and to determine whether profiling could identify drug-sensitive and drug-resistant patients by comparing subgroups for AC and TC. Results: The paper describing the chemosensitivity signatures has been withdrawn for reproducibility and data integrity problems. Consequently the scientific justification for the trial no longer exists and the trial was permanently closed on 11/1/2010. 39 patients were randomized. The overall pCR rate was 13% (95% CI, 4%-27%). The pCR rates in the genomically guided v. randomized arms were 15% (4/26) and 8% (1/13), respectively (p=0.64). The analysis by individual regimens showed: AC overall pCR rate 21% (95% CI, 6%-46%), predicted sensitive 25% v. predicted resistant 18% (p=1.0); TC overall 5% (95% CI, 0-25%), predicted sensitive 0% v. predicted resistant 9% (p=1.0). Conclusion: This underpowered analysis was conducted on the available data to examine the possibility of an adverse trend in patient outcomes. No evidence of this was seen. The data and experience from this trial can inform the design of future integral biomarker-driven trials, particularly in the preoperative setting.

BACKGROUND:

- **Genomic profiling can provide improved assessment of breast cancer biology, potentially providing improved prognostic and predictive biomarkers.**
- **Preoperative chemotherapy is a proven approach that provides an early surrogate endpoint (pCR) for validating predictive markers.**
- **In vitro/cell-line derived predictive markers were developed for doxorubicin and docetaxel and evaluated retrospectively. The translation of in vitro markers successfully could significantly accelerate biomarker development.**
- **A prospective marker-based preoperative chemotherapy trial in early-stage breast cancer with Affymetrix-based genomic profiling integral to patient treatment assignment using standard of care chemotherapy regimens was designed to validate these markers.**
- **Subsequent review of the foundational work deriving the markers and the retrospective validations showed this work to be flawed, and it has subsequently been retracted. This validation trial was therefore no longer justified and was closed to accrual.**
- **An underpowered analysis of the primary endpoint in the enrolled patients was undertaken as part of the Data Safety and Monitoring Board review and to address concerns that the flawed biomarkers might have resulted in inappropriate treatment assignments.**

PATIENTS and METHODS:

Eligibility Criteria:

- **Chemotherapy deemed appropriate by treating oncology team. No restriction on using other tests to assess chemotherapy indication (i.e. Oncotype allowed).**
- **Tumor size over 1.5 cm by clinical assessment.**
- **Any ER/PR status allowed; HER2 negative by standard clinical assessment.**
- **No T4 lesions.**
- **No contralateral cancer.**
- **Multifocal disease confined to one quadrant allowed; no multicentric disease.**

PATIENTS and METHODS (Continued):

Treatment Regimens:

- **See Figure 1 for protocol schema**
- **Standard of Care Regimens**
  - **Docetaxel 75mg/m2 and cyclophosphamide 600 mg/ m2 IV every 3 weeks for 4 cycles (TC)**

    OR
  - **Doxorubicin 60 mg/m2 and cyclophosphamide 600 mg/m2 IV every 3 weeks for 4 cycles (AC)**
- **Following completion of assigned treatment, patients were clinically re- assessed for response per standard of care with team and patient decision regarding whether to proceed to surgery or additional chemotherapy.**
- **Patients were unblinded for arm assignment and doxorubicin/docetaxel prediction scores after completing therapy.**
- **Nodal assessment was performed pre –chemotherapy per institutional standard, although not required per protocol (i.e. could also be done at definitive surgery)**

Table 1: Enrollment Demographics

(Consented and Randomized, July 2008 to July 2010)

	Consented		Randomized and Treated	
	N=57	%	N=39	%
Age				
Under 40	4	7%	2	5%
40-49	21	37%	14	36%
50-59	17	30%	13	33%
60-69	11	19%	8	21%
70-79	4	7%	2	5%
Race				
White	41	72%	29	74%
Black	13	23%	7	18%
Hispanic	1	2%	1	3%
Not Reported	2	4%	2	5%
Tumor Size				
< 4 cm	35	61%	24	62%
≥ 4 cm	21	37%	15	38%
Histology				
Ductal	45	79%	36	92%
Lobular	6	11%	3	8%
Other	6	11%	0	0%
ER Status				
ER positive	39	68%	28	72%
ER negative	16	28%	11	28%
ER unknown	2	4%	0	0%

* Withdrew	Screen Failure
3 Declined repeat biopsy	3 multicentric
1 Oncotype=11	3 inelig (PS>2; lft's; other ca)
1 Insurance concerns	1 Hybridization Failure
	4 No invasive cancer
	1 Tumor size under 1.5cm
	1 Inadequate RNA

Median tumor size is 3.2 cm

(range 1.6 cm to 7.5cm)

RESULTS:

Toxicities:

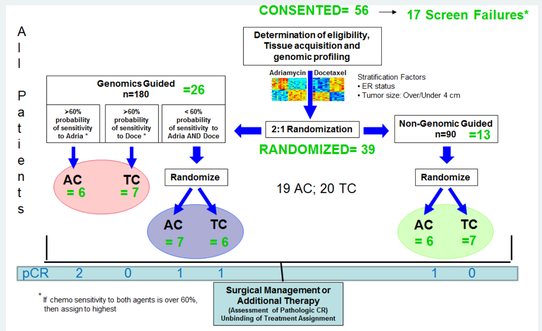
- **Consistent with those seen with standard dose AC and TC**
- **By chemotherapy:**
  - **AC: 11 of 19 with grade 3 or 4 adverse events (all neutropenia except one treatment unrelated PE)**
  - **TC: 8 of 20 with grade 3 or 4 adverse events (all neutropenia except one docetaxel reaction)**
- **By arm:**
  - **Genomically-guided: 14 of 26 with grade 3 or 4 adverse events (all hematologic except one docetaxel reaction and one treatment unrelated PE)**
  - **Control: 5 of 13 with grade 3 or 4 adverse events (all hematologic)**
  - **No treatment related deaths**

Primary Aim:

To determine in early stage breast cancer treated with primary systemic therapy whether genomic profiling for drug-sensitivity can improve the pathologic complete response rate as compared to random assignment of patients to therapy.

(pCR 26% in genomically-guided v. 13% in control, 82% power)

Figure 2: Pathologic Complete Responses (pCR) in Breast by treatment and arm (All Patients)



- **One patient randomized and assigned to genomically guided AC ineligible due to multicentric disease and treated off protocol**
- **Four patients have developed metastatic disease**
- **Two have died from recurrent disease; both had high-grade "triple-negative" disease**

Table 2: Response by final pathology and/or clinical evaluation

Overall pCR	
13% (95% CI: 4-27%)	
Genomically-guided (n=26)	Control (n=13)
15% (4/26)	8% (1/13)
P=0.64	

This analysis includes all randomized and treated patients.

Table 3: Response by final pathology only

The analysis below excludes four TC treated patients; 3 did not have definitive surgery or biopsy after completing protocol defined treatment and one had a docetaxel reaction and was treated with AC.

AC Treated Cases			
	pCR		
Genomic Prediction	No	Yes	Total
Resistant	9	2	11
Sensitive	6	2	8
Total	15	4	19
Fisher's exact p>0.99			

TC Treated Cases			
	pCR		
Genomic Prediction	No	Yes	Total
Resistant	8	1	9
Sensitive	7	0	7
Total	15	1	16
Fisher's exact p>0.99			

Conclusions

This underpowered analysis does not suggest any difference in pathologic complete response rates between the investigational marker guided arm and the control arm, or by prediction of resistance/sensitivity for either regimen. Given the conservative standard of care design of the study, there is no evidence that patients were harmed by participation. The genomic data generated is of high-quality and performed on a standard Affymetrix array. It will be made publically available for interested investigators pending completion of analyses and review.

REFERENCES

1. Potti A. et al: Genomic signatures to guide the use of chemotherapeutics. Nat Med. 12:1294 1300, 2006 (This publication has been retracted)

2. Bonnefoi, H. et.al., Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG00-01 clinical trial. Lancet Oncology, 8: 1071-1078, 2007. (This publication has been retracted)

3. http://news.sciencemag.org/scienceinsider/DukeTrialLetter.pdf

Supported by CDMRP Department of Defense Clinical Trial award grant W81XWH-07-1-0394

Generation of real-time full-genome expression data for treatment assignment in a prospective breast cancer trial

P. K. Marcom[1,5], W. T. Barry[4,5], M, J. Geradts[2], J. Olson Jr.[3,5], V. Liotcheva[5], T. Foster[5], G. Ginsburg[5], J. R. Nevins[5], M. B. Datto[2,5]

Duke Multidisciplinary Breast Program

1) Div. of Oncology, 2) Dept. of Pathology, 3) Dept. of Surgery, 4) Dept. of Biostatistics and Bioinformatics, 5) Institute for Genome Sciences and Policy
Duke University

Email: marco001@mc.duke.edu

ABSTRACT

Background: Genomic assays have been shown to provide the potential for personalized breast cancer therapy. While assays based on fixed tissues offer greater convenience, the spectrum of biology interrogated is limited. Full-transcriptome assays using microarrays are more challenging, but have the potential advantage of providing multiple prognostic and predictive signatures in one assay. We created a clinical infrastructure with the objective of obtaining full genome expression data on breast cancer samples as a clinical assay for use in a prospective trial. Methods: "Performance of Genomic Expression Profiles to Direct the Use of Preoperative Chemotherapy for Early Stage Breast Cancer" was a prospective trial analyzing genomic signatures for predicting response to doxorubicin (A) or docetaxel (T) treatment in HER2 negative cancers. Fresh-frozen cores were reviewed by the study pathologist for tumor content. RNA was then extracted and probe generated to hybridize to an Affymetrix U133Plus2.0 microarray. Microarray data quality was determined using summary metrics for U133Plus2.0 arrays and principal component analysis (PCA) plots. Results: 57 patients were consented and screened; 48 were eligible and had tissue profiling attempted. Microarray analysis was successful on 41 tumors (85%), providing data of sufficient quality to make treatment assignments. Median tumor size was 3.2 cm (range, 1.6-7.5). One sample hybridization failed QC as detected by PCA plots, two samples had insufficient RNA, and four samples had inadequate tissue. The median "study consent to treatment" and "tissue to treatment" times were 16 days (range, 10-54) and 13 days (range, 10-29), respectively. Conclusion: Real-time full-genome expression analysis on frozen tumor using an Affymetrix platform can be feasibly incorporated into an integral biomarker trial design, even in a preoperative setting. The data can potentially be analyzed for a number of prognostic and predictive signatures

BACKGROUND:

Molecular assays hold promise for personalization of cancer treatment.

- **Full-transcriptome assays:**
  - **Advantage:**
    - **Potentially validate multiple prognostic and predictive signatures using a single dataset**
    - **Dataset also available for discovery analysis**
  - **Disadvantages/Challenges**
    - **Current platforms require fresh unfixed tissue sample**
    - **Technically challenging**
- **Previous work has used in vitro derived data to define predictive genomic signatures for:**
  - **Sensitivity to common cytotoxic chemotherapy agents. (Ref. 1) (Publication has now been retracted)**
  - **Molecular pathway deregulation (Ref. 2 and 4)**
- **Prospective randomized preoperative trial in breast cancer designed to validate *in vitro* derived doxorubicin and docetaxel chemosensitivity signatures (Publication now retracted, Ref. 3)**

Clinical Trial Design:

- **Marker-guided design to test feasibility and utility of genomic information for guiding chemotherapy selection.**
- **Preoperative treatment to provide short-term clinically relevant endpoint**
- **Genomic information integral to trial conduct**
- **Choose a conservative clinical question, with importance, but reasonable clinical equipoise. For this trial, initial preoperative treatment with an anthracycline vs. a taxane based regimen.**
- **Patient Safety:**
  - **Treatment assignments that are "Standard of Care"**
  - **Flexibility in patient management**
  - **Reviewed by Duke IRB, U.S. Army Medical Research and Materiel Command (USAMRMC) Office of Research Protections (ORP) Human Research Protections Office (HRPO), and outside IRB's.**

METHODS:

Figure 1: Prospective Use of Gene Signatures

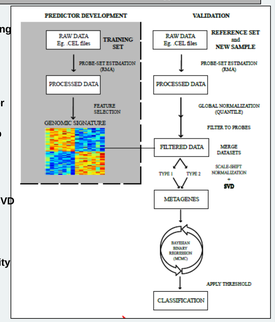
1) Development (completed prior to study):

a. **Process raw gene expression data for training set**
b. **Select features with strongest (Pearson) correlation to phenotype.**
c. **Summarize expression by a singular value decomposition (SVD) [Factors derived under Type 1 or Type 2 model]**
d. **Build a Bayesian probit regression model to predict phenotype**

2) Prospective Validation:

a. **Process investigational set normalized by SVD or scale/shift standardization.**
b. **Summarize expression values using SVD factors**
c. **Apply model to obtain a predicted probability of clinical outcome**

(Ref. 5 and 6)



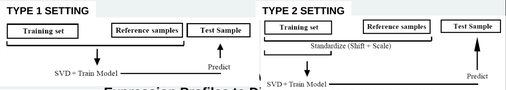
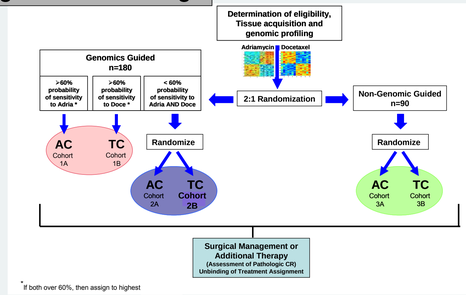


Figure 2: Trial Design

Expression Profiles to Direct the Use of Preoperative Chemotherapy for Early Stage Breast Cancer



1. **To determine in early stage breast cancer treated with primary systemic therapy whether genomic profiling for drug-sensitivity can improve the pathologic complete response rate as compared to random assignment of patients to therapy.**

   (pCR 26% in genomically-guided v. 13% in control, 82% power)

2. **To determine in early stage breast cancer treated with primary systemic therapy whether genomic profiling can identify drug-sensitive and drug-resistant patients including comparison of subgroups for the two individual regimens (AC and TC).**

RESULTS:

Figure 3: Patient eligibility and tissue acquisition

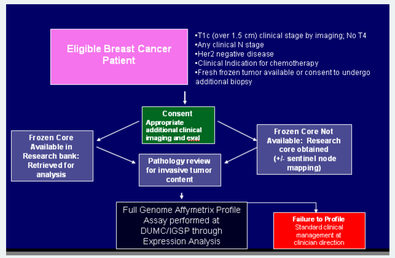


Table 1: Enrollment Demographics

(Consented and Randomized, July 2008 to July 2010)

	Consented		Randomized and Treated	
	N=57	%	N=39	%
Age				
Under 40	4	7%	2	5%
40-49	21	37%	14	36%
50-59	17	30%	13	33%
60-69	11	19%	8	21%
70-79	4	7%	2	5%
Race				
White	41	72%	29	74%
Black	13	23%	7	18%
Hispanic	1	2%	1	3%
Not Reported	2	4%	2	5%
Tumor Size				
< 4 cm	35	61%	24	62%
≥ 4 cm	21	37%	15	38%
Histology				
Ductal	45	79%	36	92%
Lobular	6	11%	3	8%
Other	6	11%	0	0%
ER Status				
ER positive	39	68%	28	72%
ER negative	16	28%	11	28%
ER unknown	2	4%	0	0%

* Withdrew	Screen Failure
3 Declined repeat biopsy	3 multicentric
1 Oncotype=11	3 inelig (PS>2; lft's; other ca)
1 Insurance concerns	1 Hybridization Failure
	4 No invasive cancer
	1 Tumor size under 1.5cm
	1 Inadequate RNA

Median tumor size is 3.2 cm

(range 1.6 cm to 7.5cm)

Figure 4: Profiling metrics

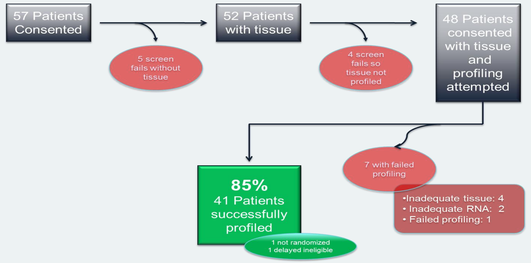
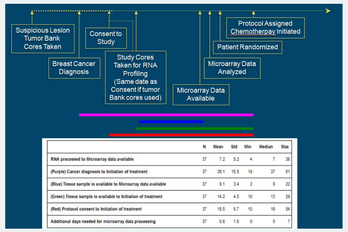


Figure 5: Profiling timelines



	N	Mean	Std	Min	Median	Max
RNA processed to Microarray data available	37	7.2	5.2	4	7	36
(Purple) Cancer diagnosis to Initiation of treatment	37	39.1	15.5	19	37	91
(Blue) Tissue sample is available to Microarray data available	37	9.1	2.4	2	9	22
(Green) Tissue sample is available to Initiation of treatment	37	14.2	4.5	10	13	29
(Red) Protocol consent to Initiation of treatment	37	16.5	6.7	10	16	54
Additional days needed for microarray data processing	37	0.6	1.6	0	0	7

Conclusions

Incorporation of real-time genomic profiling into a preoperative breast cancer treatment trial is possible, practical, and safe.

Full expression data can be generated in an appropriate time frame and be analyzed for a variety of phenotypic signatures. While candidate chemosensitivity signatures were examined in this trial, the expression data can also be used for discovery of new phenotypic signatures.

Statistical methodologies for applying *in vitro* derived genomic signatures to "one at a time" clinical samples were developed and implemented successfully.

HOWEVER, the scientific work describing and validating the in vitro derived signatures was flawed and the supporting publications have been retracted. Since this work provided the justification for conducting the trial, this study has now been closed.

REFERENCES

1. Potti A. et al: Genomic signatures to guide the use of chemotherapeutics. Nat Med. 12:1294 1300, 2006 (This publication has been retracted per discussion above)

2. Bild AH et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 439:353-7, 2006

3. Bonnefoi, H. et.al., Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG00-01 clinical trial. Lancet Oncology, 8: 1071-1078, 2007. (This publication has been retracted)

4. West M et al: Predicting the clinical status of human breast cancer by using gene expression profiles. PNAS 98: 11462-7, 2001

5. Barry W. et. al: Utilization of genomic signatures for chemotherapy response in prospective clinical studies, ASCO Meeting Abstracts Jun 14, 2010: 10513

6. Barry WT et al: Intra-tumor heterogeneity and precision of microarray-based predictors of breast cancer biology and clinical outcome. J Clin Oncol. 28(13), 2198-206, 2010

7.. Lee JK et al: Prospective comparison of clinical and genomic multivariate predictors of response to neoadjuvant chemotherapy in breast cancer. Clin Can Res 2010; 16: 711.

Supported by CDMRP Department of Defense Clinical Trial award grant W81XWH-07-1-0394

05 _Kuderer_ A cross-platform comparison... 2012-11-30_1642 .pdf

A Cross-Platform Comparison of Genomic Signatures and OncotypeDx Score: Discovery of Potential Prognostic/Predictive Genes and Pathways

Nicole M Kuderer[1]; William T Barry[1]; Joseph Geradts; Geoffrey S Ginsburg; Gary H Lyman, Michael Datto, Vlayka Liotcheva; Pam Isner; Timothy Veldman; Pankaj Agarwal; Shelley Hwang; Neal Ready; Paul K Marcom

[1]Co-First Author, Duke Cancer Institute; Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute; Duke Center for Personalized Medicine, Institute for Genome Sciences and Policy; Comparative Effectiveness and Health Outcomes Research, Duke University Medical Center, Durham, NC

BACKGROUND

- Microarray assessment of breast cancer demonstrates disease subsets among the major breast cancer biologic categories (ER, PR, HER-2).
- These have likely additional prognostic and treatment (predictive) implications.
- Additional prognostic and predictive biomarkers and optimization of existing genomic platforms are needed to improve personalized breast cancer care.

OBJECTIVES

- The objective of our study in ER+ breast cancer patients is:
  - To perform a cross-platform comparison of existing prognostic genomic signatures and the true OncotypeDx recurrence score
  - To perform discovery of additional prognostic / predictive genes or pathways in ER+ breast cancer patients

METHODS

Sample Selection and Patient Characteristics

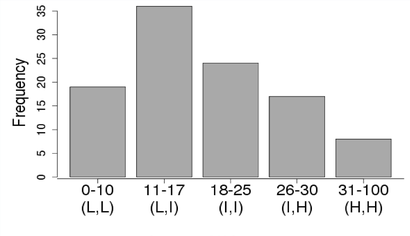
- Selection of early-stage ER+ patients with adequate RNA from fresh frozen tumor specimens and a concurrent 21-gene OncotypeDX recurrence score (RS) were identified across independent Duke studies linked to routine prospective breast biospecimen collection.
- After quality control of the microarray data, and normalization for technical batches in the laboratory procedures, data are available for a total of 104 ER+ patient samples annotated with Oncotype RS and other patient and tumor characteristics.
- Originally 73 patient samples were eligible. Since abstract submission an additional 31 ER+ breast cancer patients were found eligible and are included in this expanded analysis.
- Patient Characteristics:
  - 104 ER-positive early-stage breast cancer patients
  - Median age 55 (range 35 – 86)
  - Postmenopausal 63%
  - Caucasian 87%, African American 11%

Data Analysis

- Expression estimates for Affymetrix H133 Plus 2.0 microarrays were reviewed for quality control (Owzar et al.) and normalized across batches with ComBat (Johnson et al).
- The association between mRNA expression by Affymetrix to OncotpeDX test was investigated at the feature-level using linear models for microarrays (LIMMA, Smyth et al.) and the Benjamini-Hochberg method was employed to account for multiple testing.
- Affymetrix-based models of prognostic breast cancer signatures, including 21-gene model in Oncotype, the 70-gene model in MammaPrint, the 50-gene model in PAM50, and the GENIUS model were computed using algorithms from Fan et al and Haibe-Kains et al. Association of microarray model scores to true Oncotype recurrence score was assessed using Spearman rank correlation, and gene sets to the models were tested using SAFE (Barry et al).
- 22 key pathway-signatures are built into the ScoreSignature module (Chang et al) hosted on the Duke Institute for Genome Sciences and Policy (IGSP) instance of GenePattern. The association of predicted probabilities from each pathway signature with true Oncotype recurrence score is assessed by Spearman rank correlation with corrections for multiple testing using the Benjamini-Hochberg method. Agglomerative hierarchical clustering of patient samples and pathways is provided by the ScoreSignature module, and results are displayed as a heat map of predicted probabilities.
- A shrinkage-based regression approach (Hastie) was used to build a new microarray model of Oncotype score from the Affymetrix microarray data. Using a 2:1 random split into training and test patient cohorts, a 14-gene signature was derived for the linear OncotypeDX score in the training set. After locking the Oncotype signature, it was independently assessed in the test cohort.

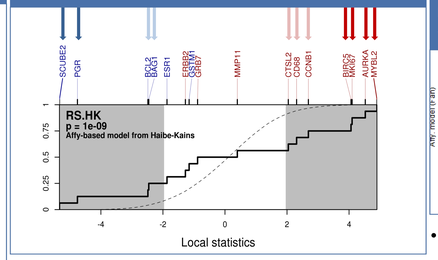
RESULTS

Figure 1 Distribution of OncotypeDx Recurrence Scores (RS)



- **Figure 1**: Distribution of OncotypeDX scores for our ER+ breast cancer patients by low (L), intermediate (I), and high (H) risk OncotypeDX categories per Paik et al (intermediate scores 18-30) and per TAILORx trial (intermediate scores 11-25).

Figure 2 Association of Microarray Oncotype Genes to true Oncotype Score



- **Figure 2**: General association between the microarray oncotype genes and true OncotypeDX score was highly significant by SAFE (Barry et al). In the SAFE plot, the direction of the microarray differential expression of oncotype probe sets largely agrees with the model reported for the Oncotype DX assay - with exceptions likely attributable to our cohort being predominantly ER+/HER2-.

Table 1 Discovery of Prognostic Genes associated with Oncotype Recurrence Score (RS)

Rank	Affymetrix ID	Fold change (10 units RS)	Adjusted p-value	Gene	Full Gene Name	Gene Functions
1	206632_s_at	1.54	4.21E-05	APOBEC3B	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B	Proteins may be RNA editing enzymes and have roles in cell growth or cell cycle control and immune response.
2	219197_s_at	0.53	0.00025	SCUBE2	Signal peptide, CUB and EGF-like domain-containing protein 2	Suspected breast cancer suppressor protein in part by suppression of β-catenin pathway
3	1552368_at	1.20	0.00033	CTCFL	CCCTC-binding factor (zinc finger protein)-like	Involved in gene regulation, utilizes different zinc fingers to bind varying DNA target sites.
4	204002_s_at	0.86	0.00041	ICA1	Diabetes mellitus type I autoantigen	Role in regulating the early transport of insulin secretory granule proteins
5	211211_x_at	1.16	0.00041	SH2D1A	T cell signal transduction molecule SAP	Protein that plays a major role in the bidirectional stimulation of T and B cells
6	242326_at	1.13	0.00082	COL22A1	Collagen, type XXII, alpha 1	Cell adhesion ligand for skin epithelial cells and fibroblasts
7	225728_at	0.66	0.00096	SORBS2	Arg binding protein 2	Family of adaptor proteins believed to play roles in cell adhesion, cytoskeletal organization, and signaling
8	207746_at	1.28	0.00096	POLQ	Polymerase (DNA directed), theta	Potential role in base excision repair (BER) and chemotherapy resistance
9	219836_at	1.27	0.00096	ZBED2	Zinc finger BED domain-containing protein	Unknown function
10	201710_at	1.20	0.00100	MYBL2	V-myb myeloblastosis viral oncogene homolog (avian)-like 2	Possesses both activator and repressor cell cycle activities. Activates the cell division cycle 2, cyclin D1, and insulin-like growth factor-binding protein 5 genes.
14	228554_at	0.53	0.00172	PGR	Progesterone receptor	Mediates the physiological effects of progesterone
32	208079_s_at	1.34	0.00303	AURKA	Aurora kinase A	Cell cycle kinase involved in microtubule formation and/or stabilization. Likely role in tumor development and progression.
57	208368_s_at	1.14	0.00459	BRCA2	Breast cancer 2, early onset	Homologous recombination pathway for double-strand DNA repair
87	212021_s_at	1.25	0.00611	MKI67	Antigen identified by monoclonal antibody Ki-67	Cell proliferation
90	224753_at	1.22	0.00633	CDC45	Cell division cycle associated 5	Cell cycle and ERK pathway
142	202095_s_at	1.34	0.00799	BIRC5	Apoptosis inhibitor 4 / baculoviral IAP repeat containing 5	Proteins prevent apoptotic cell death

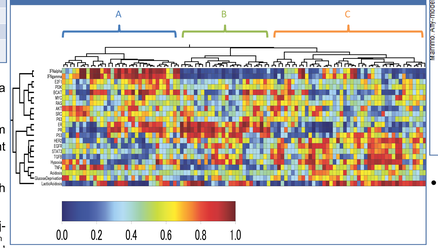
- The top 194 Affymetrix probe sets from the Affymetrix platform had a FDR <1%.
- Shown are (a) all genes with adjusted $p < 0.001$, (b) genes from OncotypeDX that reached statistical significance, and (c) other relevant genes linked to breast cancer that were statistically significant.
- Several known cancer-associated genes are highly correlated with prognosis:
  - some are on the DX platform, e.g.: SCUBE2 (ranked 2nd, adj-p=0.002), MYBL2 (ranked 10th, adj-p=0.001) and PGR (ranked 14th, adj-p=0.002);
  - and some are not part of the DX platform, e.g.: BRCA2 (ranked 57th, adj-p=0.005), and CDCA5 (ranked 90th, adj-p=0.006).

Table 2 Association of key oncogenic Pathway-Signatures to Oncotype Score

Rank	Path	Corr	P.val	Adj.p
1	P53	-0.45	1.34E-06	2.94E-05
2	PR	-0.31	0.0012	0.0129
3	BCAT	0.25	0.0112	0.0766
4	RAS	0.23	0.0172	0.0766
5	SRC	0.23	0.0174	0.0766

- **Table 2**: For association with prognostic score, two oncogenic pathways pass a FDR of 5% (light blue), and three additional pathways pass a less stringent FDR threshold of 10% (darker blue).

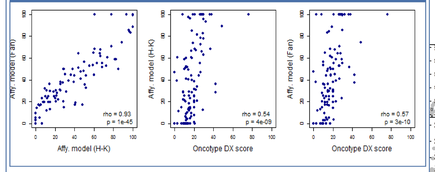
Figure 3 Patient Clustering by Pathway Signatures



- **Figure 3**: In unsupervised hierarchical clustering, pathway signatures group patients into three major genomic clusters (A, B, C).
- The potential clinical significance of these pathway-derived groups and connections to luminal-type biology will be further explored.

Figure 4 Association of Microarray-Oncotype Signatures with true OncotypeDX Score:

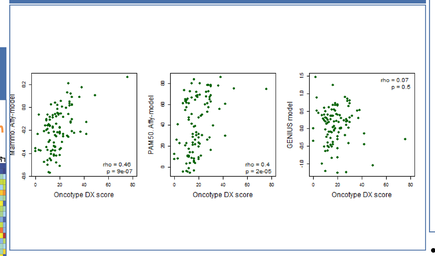
Microarray-Oncotype specific models using the 21 Oncotype genes: via Haibe-Kains et al and via Fan/Perou et al



- **Figure 4**: Two existing Affymetrix-based microarray models for Oncotype score from Fan/Perou and Haibe-Kains differ in how they pick the Affymetrix probe sets to represent each oncotype gene, and also differ in their respective algorithms.
- Nevertheless, output from these two microarray-oncotype signatures are highly correlated with each other (left panel, rho=0.93).
- Despite the highly significant correlation of the microarray gene sets and Oncotype scores, agreement of both microarray-oncotype signatures with true Oncotype score is relatively poor with a prominent upward-shift in both microarray models (middle and right panels).

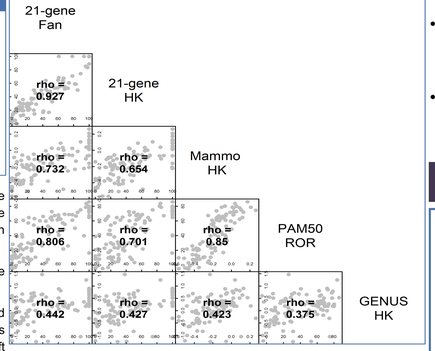
Figure 5 Association of Other Breast Cancer-Specific Prognostic Microarray Signatures and true Oncotype Score:

PAM50 w/ risk-of-relapse and proliferation scores (Parker/Perou et al)
70-gene signature for MammaPrint (Van't Veer, via Haibe-Kains et al)
GENIUS (Haibe-Kains/Quackenbush et al)



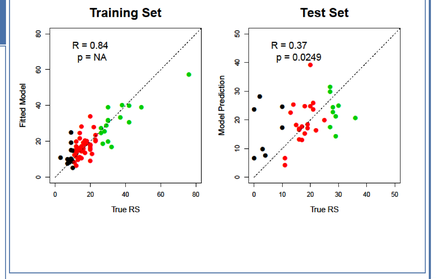
- **Figure 5**: The relationship of Affymetrix-based models for other breast cancer-specific prognostic models to true Oncotype DX score is lower than the microarray-oncotype model, although statistically significant correlation is still seen with all except the GENIUS prognostic model (right panel, rho=0.07, p=0.5).

Figure 6 "Scatter Plot Matrix" for Relatedness of the different Microarray-based Breast Prognostic Models



- **Figure 6**: The following "scatter plot matrix" displays the relatedness of the various microarray models of breast cancer prognosis.
- Microarray models show good correlation among each other for ER+ patients, with the exception of the GENIUS model.

Figure 7 New Oncotype-Microarray Model Development and Model Testing for true OncotypeDX recurrence score (RS)



- **Figure 7**: A shrinkage-based regression approach was used to build a new microarray model for the true OncotypeDX recurrence score, based on a 2:1 random split of the overall patient cohort into a training set (n=68) and a test set (n=36).
- By 10-fold cross validation in the training set, a shrinkage factor to the LASSO model identified a 14-gene microarray prediction model for the continuous OncotypeDX score.
- Our microarray-oncotype model was independently applied to the test set, resulting in a positive level of association with the true OncotypeDX score (Pearson r=0.37, p=0.025), but without sufficient accuracy yet for Oncotype prediction for clinical use.

DISCUSSION

- Even in ER+ breast cancer patients, OncotypeDX provides limited treatment guidance for intermediate risk patients. Large proportions of breast cancer patients are currently classified at intermediate risk per OncotypeDX recurrence score. Additional prognostic / predictive biomarkers could be informative for improved personalized chemotherapy and treatment decisions.
- Previously reported microarray-based surrogate Oncotype signatures are generally concordant. However, they have suboptimal correlation with the true Oncotype score. They are limited by sample size and the availability of true OncotypeDX scores.
- Our findings are from largely ER+/HER- patients with available OncotypeDX score, but are still limited by the cohort size. Further assessment, validation, and reproducibility of our findings are needed prior to clinical use.

CONCLUSIONS

- **Previously reported microarray-based surrogate Oncotype signatures are limited by sample size and the availability of the actual OncotypeDX recurrence scores, which are necessary for a precise microarray-based Oncotype signature development.**
- **In ER+ breast cancer, we have identified additional key cancer-associated genes, pathways, and genomic signatures associated with breast cancer prognosis, potentially providing insights into treatment opportunities for ER+ breast cancer.**
- **Further validation of these additional prognostic genes and pathways hold the promise to improve personalized breast cancer care in ER+ breast cancer patients, especially in intermediate risk OncotypeDX patients.**
- **Cohort(s) with concurrent OncotypeDX score and recurrence-free survival data are required to prove the clinical validity and utility of new prognostic biomarkers in ER+ early stage breast cancer patients.**

References:
(**1**) Barry, W.T., Ann Appl Stat. 2008; 286-315. (**2**) Smyth G.K., Stat Appl Genet Mol Biol. 2004; 3:Article3. (**3**) Johnson W.E. et al., Biostatistics. 2007; 8:118-27. (**4**) Hastie, T., Springer. 2009; xxii, 745. (**5**) Fan C. et al., N Engl J Med. 2006; 355:560-9. (**6**) Haibe-Kains B. et al., BMC Genomics. 2008; 9:394. (**7**) Chang J.T. et al., BMC Bioinformatics. 2011; 12:443. (**8**) Gatza M.L. et al., Proc Natl Acad Sci U S A. 2010; 107:6994-9. (**9**) Prat A. et al., Ann Oncol. 2012; 23:2866-73. (**10**) Owzar K. et al., Clin Cancer Res. 2008; 14:5959-66. (**11**) Paik S. et al., N Engl J Med. 2004; 351:2817-26. (**12**) Parker J.S. et al., J Clin Oncol. 2009; 27:1160-7. (**13**) van 't Veer L.J. et al., Nature. 2002; 415:530-6.

FUNDING / ACKNOWLEDGEMENTS

NCI: RC2CA14041-01; W81XWH-07-1-0394; NCI: K-Award (NK);
ASCO Young Investigator Award (NK); Supported by the CJL Foundation (WB)

Presented at 2012 AACR SABCS Symposium, San Antonio, TX, December 4-8, 2012

Retrospective evaluation of precision of gene-expression-based signatures of prognosis and tumor biology in replicate surgical biospecimen from patients with breast cancer

William T. Barry[1], P. Kelly Marcom[2], Joseph Geradts[2], Michael B. Datto[2]

[1]Dana-Farber Cancer Institute, Boston MA [2]Duke University Medical Center, Durham, NC

Abstract

Background: Numerous gene-expression signatures have been developed for breast cancer. However, assessments of their validity continue to largely ignore the impact of intratumor heterogeneity and technical variation on clinical utility. Here, the collection of replicate specimen in Barry et al (2010) is used to evaluate a broad collection of gene-expression signatures of prognosis and tumor biology.

Methods: 18 patients with multiple frozen cores (one patient with quadruplicate, twelve with triplicate, and five with doublet samples) were previously identified in the Duke Breast SPORE tissue repository. Cores were assessed for percent invasive cancer cellularity, and tumor size, grade, and ER/PR status. RNA was extracted and hybridized to AffymetrixH133Plus2.0 microarrays. Expression signatures of prognosis and tumor biology were developed or translated to the Affymetrix platform using routines by Prat et al. (2012) and Haibe-Kains et al. (2012). Association between signatures, and precision among replicates are evaluated using Pearson and intraclass correlation. Coefficients are reported with 95% confidence intervals.

Results: Among prognostic signatures, an imputed 70-gene signature of MammaPrint had the highest level of precision (ICC=0.96, 0.91-0.98); strong concordance is seen with Affymetrix-based PAM50 risk-of-relapse (ICC=0.82, 0.65-0.92); 16 of 18 patients had constant subtype calls. Substantially lower concordance was seen in the GENIUS prognostic model (ICC=0.62, 0.35-0.82). In ER+ patients (n=13), two algorithms to impute the 21-gene model of OncotypeDX recurrence score were concordant (r =0.98, ICC= 0.90 and 0.83) and distinct from the GENIUS ER+ score (average r=0.74, ICC=0.78). Imputed models for the Rotterdam 76-gene model (+ER status), GGI (+grade), and ROR-T (+tumor size) showed lower levels of correlation (0.47<rho<0.64), suggesting independent prognostic information is conveyed by the unique combinations of clinical and genomic factors. High correlation is seen between models of ER status ranging from single-gene data (ICC=0.90) to signatures from archived specimen (ICC=0.96) and cell lines (ICC=0.85), with almost complete agreement to IHC results. Two signatures of PI3K from specimen with mutation data were partially correlated and concordant (r=0.52, ICC= 0.96 and 0.91), while two signatures from overexpression systems were uncorrelated, and one with poor precision (r=0.12, ICC= 0.92 and 0.46).

Conclusions: The number of genomic signatures in breast cancer from archived specimen and cell-line experiments continues to grow, but there are limited resources for validating their prognostic/predictive value in patient populations. Reproducibility across biological replicates is a critical component in establishing clinical utility of a signature that is distinct from using technical replicates for the repeatability of analytes on the array platform. We demonstrate how archived specimen can confirm reproducibility in the 'Test Validation Phase' of biomarker development, as advocated by the Institute of Medicine, and inform trial designs to prospectively test clinical utility.

Update: Additional pathway signatures of PR, p53 and EGFR are reported. Some of the findings are expanded to a second cohort of 70 replicate samples from 24 HER2- patients. Replicate pre- and post- treatment cores available from the clinical trial "Utilization of Genomic Signatures to Direct Use of Primary Chemotherapy in Early Stage Breast Cancer" (NCT00636441) were analyzed using the same procedures employed for Cohort 1 samples. The impact of tumor cellularity on the precision and accuracy of signatures is calculated using standard deviation .

Number of repl.	Cohort 1	Cohort 2
2	5	8
3	12	10
4	1	16
Total samples	50	70

Results

Signature/algorithm sources:

- R/Bioconducor package, `genefu`, by Haibe-Kains et al.:
  70-gene (cf. Mammoprint); 16-gene (cf. Oncotype); 76-gene (cf. Rotterdam); GGI; PIK3CA; GENIUS
- R source code from the Perou laboratory (http://peroulab.med.unc.edu/):
  21-gene (cf. Oncotype); 50-gene PAM50 intrinsic subtype and risk-of-relapse (ROR); PI3Kwt and PI3Kmut (Hutti).
- GenePattern module, ScoreSignature, by Gatza/Chang for genomic signature of oncogenic pathways (OncPath) from a variety of sources:
  AKT, Beta-catenin, E2F1, EGFR, ER, HER2, IFNα, IFNγ, MYC, p53, PI3K, PgR, SRC, STAT3, TGFβ, p63, RAS, TNFα
- R source code from Barry et al. (2010, in prep.): ER, p53 (Miller, 2005)
- Single-probesets identified by Gong et al. (2006).

Table 1: Agreement within cohort 1	ICC (95% CI)
Prognostic	
70-gene (Mammoprint)	0.96 (0.91, 0.98)
PAM50 ROR	0.82 (0.65, 0.92)
GENIUS	0.57 (0.29, 0.79)
21-gene and 16-gene (Oncotype DX) in ER+ patients.	0.90 (0.77, 0.97) 0.83 (0.64, 0.94)
GGI, 76-gene (Rottingham); ROR-T	-NA-
Predictive of tumor biology	
ER: IHC in archived spec. (Barry, GSE3494)	0.97 (0.93, 0.99)
ER: single-probeset (Gong)	0.93 (0.86, 0.97)
ER: in vitro cell-lines (Gatza)	0.85 (0.70, 0.94)
PR: single-probeset (new)	0.90 (0.79, 0.96)
PR: mut in archived spec. (Gatza, GSE3494)	0.90 (0.80, 0.96)
p53: mut in archived spec. (Miller, GSE3494)	18 of 18 concordant
p53: mut in archived spec. (Gatza, GSE3494)	0.92 (0.84, 0.97)
PI3K: PIK3CA mut. in archived spec. (Loi)	0.95 (0.89, 0.98)
PI3K: in vitro overexpression of mut. (Hutti)	0.91 (0.82, 0.96)
PI3K: in vitro overexpression of wt. (Hutti)	0.90 (0.80, 0.96)
PI3K: in vitro overexpression (Gatza)	0.46 (0.16, 0.73)
EGFR: single-probeset (new)	0.88 (0.79, 0.95)
EGFR: in vitro overexpression (Gatza)	0.44 (0.14, 0.71)

Figure 1: Agreement of gene-signatures for tumor biology with IHC (cohort 1)

ER		Single gene (>9.6, pre-spec.)			Archived spec. (>0.63, opt.)			OncPath (>0.14, opt.)		
		-	-/+	+	-	-/+	+	-	-/+	+
IHC -	5	5	0	0	5	0	0	4	1	0
IHC+	13	1	0	12	1	0	12	0	1	12

As reported in Barry 2010, very strong agreement is seen with IHC, and few discordant calls (labeled '-/+') are observed among replicates when using either a pre-specified threshold (Gong et al), or cut-points to continuous scores which maximize concordance among replicates.

PR		Single gene (>4.0, opt.)			OncPath (>0.19, opt.)		
		-	-/+	+	-	-/+	+
IHC -	6	1	4	1	5	1	0
IHC+	12	0	3	9	1	1	10

For the progesterone receptor, the gene signature from archived specimen (Gatza) outperforms any single-gene measurement on the Affymetrix array that is optimized for precision/accuracy.

EGFR		Single gene (>8.5, opt.)			OncPath (>0.5)		
		-	-/+	+	-	-/+	+
IHC -	12	9	3	0	3	4	5
IHC+	6	0	0	6	1	5	0

Conversely, for EGFR, a single-gene measure closely correlates with IHC and has good concordance (15 out of 18); while the OncPath signature from an overexpression system performs poorly.

Figure 2: Scatterplots of agreement within prognostic signatures

In Cohort 1 (n = 18 patients), the genomic prognostic signatures (a) 70-gene, (b) PAM50-ROR, and (c) GENIUS show varying levels of agreement among replicates (Table 1). In each case, the five patients that were ER- by IHC were classified as having worse prognosis on average in the replicate Shown as horizontal lines are pre-defined thresholds to the 70-gene (-0.3) and PAM50-ROR (22, low versus intermediate or high). Thresholds are observed to also differentiate the ER+ patients with minimal discordance (2 and 3 of 13 ER+ cases, respectively). For the GENIUS model, no threshold is provided in the algorithm. Regardless, a higher discordance rate is seen for any threshold between -1 and 1 that differentiates the ER+ subjects.

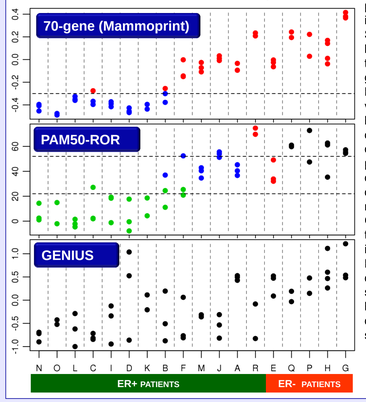
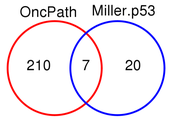
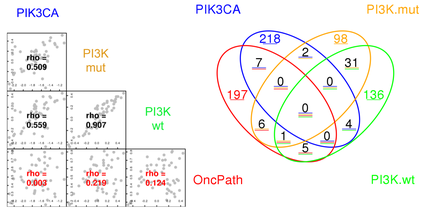


Figure 3: Correlation among signatures of tumor biology

p53 (GSE3494)		OncPath (>0.71 *opt.)		
		-	-/+	+
Miller et al.	-	22	2	0
	-/+	1	5	0
	+	1	0	11

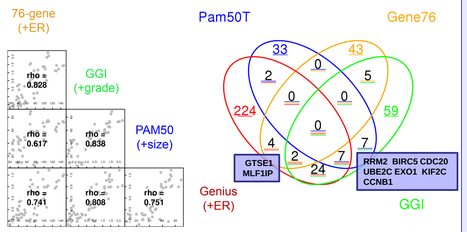


Two multi-gene signatures of p53 mutation status were separately developed using the same set of archived specimen. Despite minimal overlap of gene-sets and different set sizes in each model, strong agreement and precision is seen in predictions.



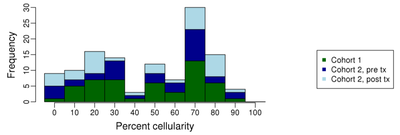
Among four PI3K genomic signatures: (a) Strong agreement is seen in predictions based on the two signatures from Hutti, despite distinct forms of PI3K being transfected. Moderate agreement is seen with the PIK3CA signature derived from archived specimen despite only 4+2 overlapping genes. The OncPath signature by Gatza et al. has no agreement with the other signatures, and poor precision among replicates (Table 1.)

Figure 4: Correlation among prognostic clinical+genomic signatures

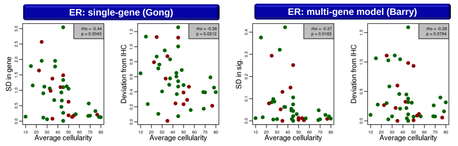


There is of minimal overlap in genes and clinical factors among the 4 signatures (9 genes in 3 or more). In spite of this, strong positive correlation is seen in Cohort 1.

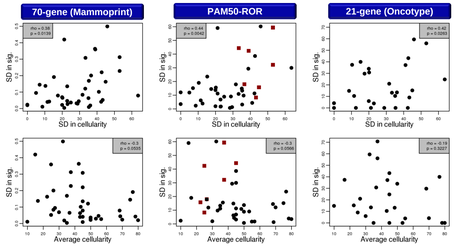
Figure 5: Impact of tumor cellularity on precision and accuracy



Combining Cohorts 1 and 2, tumor cellularity assessed by histopathology ranges from 0 – 90%. Impact of cellularity on precision of signatures was measured by standard deviation (SD). The panels below for two ER signatures are consistent with other signatures. Also, cellularity is also shown decrease accuracy relative to ER status by IHC pretreatment.



For 70-gene, 50-gene and 21-gene prognostic signatures, all lose precision when (a) average cellularity is low, and (b) variability in cellularity is high. Further, for the PAM50 assay, several patient samples with low cellularity are flagged as 'Normal-like' (red sq.)



Summary

The clinical utility of any genomic signature requires a strong degree of precision across biological replicates of tissue/blood sample. However, this remains underexplored in most retrospective and prospective validation studies. Here we use replicates from an initial cohort of 18, and expanded cohort of 24 patients to evaluate a series of genomic signatures for breast cancer that have been made publically available.

- Affymetrix-versions of the 70-gene (Mammoprint) and 50-gene (PAM50) show substantial agreement by intraclass correlation, and consistently segregate ER+ tumors into low vs. intermediate risk.
- Less precision is seen with the GENIUS model.
- Prognostic models that combine genomic and clinical information are highly positively correlated with one another, despite largely non-overlapping factors to the models.
- Less correlation is seen among independently derived signatures of PI3K
- In general, tumor biology signatures derived from candidate genes or profiling archived tumors show greater precision than signatures derived from in vitro over-expression experiment.
- Cellularity is shown to impact precision of signatures of tumor biology, with notable increases in variance in some samples with <45% cellularity. This is further shown to increase deviation from the pre-treatment ER status by IHC.
- Cellularity impacts precision of prognostic signatures to a similar degree. A larger cohorts with outcome information and varying cellularity would be needed to evaluate the extent to which prognostic value is lost.

References

Barry WT. A novel method for testing the association of high-dimensional genomic signatures to clinical outcomes. In preparation.

Barry WT, Kernagis DN, et al. Intratumor Heterogeneity and Precision of Microarray-Based Predictors of Breast Cancer Biology and Clinical Outcome. J Clin Oncol. 2010;28(13):2198-206.

Gong Y, Yan K, et al. Determination of oestrogen-receptor status and ERBB2 status of breast carcinoma: a gene-expression profiling study. Lancet Oncol. 2007;8(3):203-11.

Haibe-Kains B, Desmedt C, et al. Comparison of prognostic gene expression signatures for breast cancer. BMC Genomics. 2008;9:394.

Hutti JE, Pfefferle AD, et al. Oncogenic PI3K mutations lead to NF-kappaB-dependent cytokine expression following growth factor deprivation. Cancer Res. 2012;72(13):3260-9.

Loi S, Haibe-Kains B, et al. PIK3CA mutations associated with gene signature of low mTORC1 signaling and better outcomes in estrogen receptor-positive breast cancer. PNAS 2010;107(22):10208-13.

Miller LD, Smeds J, et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. Proc Natl Acad Sci U S A. 2005;102(38):13550-5.

Parker JS, Mullins M, et al. Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. J Clin Oncol. 2009;27(8):1160-7.

Acknowledgements

Supported by Duke Breast Specialized Program of Research Excellence (SPORE) Grant No. 5P50CA068438, and Congressionally Directed Medical Research Programs Department of Defense Clinical Trial Grant No. W81XWH-07-1-0394 (P.K.M.). This work was also funded in part by the CJL Foundation.

Printed by Call4Posters