Gene expression profiling and breast cancer care: What are the potential benefits and policy implications?

Nina Oestreicher, PhD MS,^{1,2} Scott D. Ramsey, MD PhD,^{1,2} Hannah M. Linden, MD,^{1,2} Jeannine S. McCune, PharmD,^{1,2} Laura J. van't Veer, PhD,³ Wylie Burke, MD PhD,¹ and David L. Veenstra, PhD PharmD,¹

Purpose: Gene expression profiling has been proposed as an alternative to clinical guidelines to identify high-risk patients for adjuvant chemotherapy. However, the outcomes associated with gene expression profiling are not clear, and guidelines for the appropriate use of genomic technologies have not been established. Methods: We developed a decision analytic model to evaluate the incremental cost and quality-adjusted life years of gene expression profiling versus NIH clinical guidelines in a hypothetical cohort of premenopausal early stage breast cancer patients 44 years of age. We conducted empirical analyses and identified literature-based data to inform the model, and performed probabilistic sensitivity analyses to evaluate uncertainty in the results. We interpreted the implications of our findings for treatment guidelines and policies. Results: Use of gene expression profiling resulted in an absolute 5% decrease in the proportion of cases of distant recurrence prevented, 0.21 fewer quality-adjusted life years, and a cost savings of \$2882. The chosen test cutoff value to identify a tumor as poor prognosis and the cost of adjuvant chemotherapy were the most influential parameters in the analysis, but our findings did not change substantially in sensitivity analyses. Regardless of the test cutoff used to identify a poor prognosis tumor, the gene expression profiling assay studied in our analysis, at its current level of performance, did not attain the threshold sensitivity (95%) necessary to produce equal or greater quality-adjusted life years than NIH guidelines. Conclusion: Although the use of gene expression profiling in breast cancer care holds great promise, our analysis suggests additional refinement and validation are needed before use in clinical practice. Genet Med 2005:7(6):380-389.

Key Words: gene expression profiling, microarray analysis, genetic tests, breast cancer, cost effectiveness

Breast cancer is the leading incident cancer among women of all major ethnicities in the United States and is the second highest source of cancer mortality.¹ Adjuvant chemotherapy has been shown to increase recurrence-free and overall survival,² but also may produce significant toxicity in the patient. Chemotherapy may cause short-term complications such as alopecia, nausea/vomiting, and myelosuppression and may lead to longer term complications such as permanent ovarian failure in premenopausal patients.^{3,4} Current NIH clinical guidelines⁵ recommend adjuvant chemotherapy for women with tumors larger than 1 cm or lymph node involvement. Additionally, tumor markers such as HER2 and histologic

Received: January 10, 2005.

Accepted: April 4, 2005.

DOI: 10.1097/01.GIM.0000170776.31248.75

grade are used for risk assessment.^{6,7} Despite widespread use, these criteria are imprecise predictors of distant recurrence.⁸

Gene expression profiling (GEP) utilizing DNA microarrays9 or RT-PCR10 has been proposed as an alternative approach to identify patients for adjuvant chemotherapy,11 potentially sparing low-risk patients from this treatment. There are predominantly two gene expression profiles currently marketed for clinical use in breast cancer. One of these profiles, MammaPrint, was developed by van't Veer and colleagues at the Netherlands Cancer Institute and is marketed in Europe (and soon will be marketed in the United States) by Agendia.¹² The other test, Oncotype DX, is marketed by Genomic Health, Inc. in the United States. The assay marketed by Agendia utilizes a 70-gene microarray-based profile performed on fresh frozen tissue and is intended for patients younger than 55 years with Stage I invasive breast cancer or Stage II node negative invasive breast cancer. In contrast, the test marketed by Genomic Health uses a 21-gene profile utilizing RT-PCR for expression analysis on paraffin-embedded tissue and is intended for patients with node-negative, estrogen receptorpositive (ER+) disease who are taking tamoxifen. Both com-

From ¹University of Washington, Seattle, Washington; ²Fred Hutchinson Cancer Research Center, Seattle, Washington; ³The Netherlands Cancer Institute, Amsterdam, The Netherlands.

Nina Oestreicher, PhD, MS, Kaiser Permanente Division of Research, 2000 Broadway, Oakland, CA 94612.

panies are conducting clinical trials in the United States and Europe to validate their respective expression profiles. Also, there are gene expression profiles in development, such as the one by Ipsogen, a biotechnology company based in France.¹³ In considering the adoption of GEP in clinical practice, a quantitative evaluation of the clinical and economic outcomes and impact on patient quality of life can clarify the potential tradeoffs compared to current practice. Despite the potential importance of this type of analysis to clinical decision-making about GEP, there have been few such evaluations.¹⁴

The objective of this study was to estimate the cost-effectiveness of the Netherlands Cancer Institute GEP assay versus NIH guidelines for the identification of early stage breast cancer patients who would benefit from adjuvant chemotherapy, and to assess the implications of our findings for practice guidelines. The profile we used for our analysis has shown promise for accurate prediction of high-risk women (compared to NIH guidelines) in early studies.^{9,15} We directly evaluated test characteristics in our analysis to assess the accuracy and prognostic value of identifying high-risk women using GEP.

MATERIALS AND METHODS

We utilized decision analytic techniques to compare GEP and NIH clinical guidelines as prognostic tests to evaluate the risk of distant recurrence for women with early stage breast cancer. Our analysis considered a hypothetical cohort of premenopausal women averaging 44 years of age newly diagnosed with Stage I/II breast cancer. The demographic and clinical characteristics (e.g., age and nodal status) of this target population were chosen to be similar to those of the Netherlands Cancer Institute cohort, upon which our estimates of test performance (sensitivity, specificity, and positive predictive value) are based, and for which we derived empirical estimates for some model parameters.9 Briefly, the Netherlands Cancer Institute cohort comprised a consecutive series of 295 breast cancer patients 52 years or younger who had mastectomy or breast conserving surgery and radiotherapy (if indicated). Fifty-one percent of the cohort had lymph node-positive disease, 77% had ER+ disease, 47% had tumors larger than 2 cm, and 40% had high grade tumors.

The initial component of our model consisted of a decision tree (Fig. 1) that modeled the prognostic categorization and treatment of women with early stage breast cancer during the 6-month period following breast cancer diagnosis. At the time of diagnosis, the use of either prognostic test led to identification of women at high risk (poor prognosis) or at low risk (good prognosis). The results of GEP are a continuous measure, which was dichotomized as "good prognosis" or "poor prognosis" based on a designated test cutoff from the first validation study of GEP15 that yielded a 10% false-positive rate. The gold standard for the estimation of test performance was the actual distant recurrence events in the Netherlands Cancer Institute cohort. The outcomes of correctly classified cases and misclassified cases (i.e., those women with a poor prognosis tumor who were classified as good prognosis and vice versa) were taken into account in the model. For example, misclassified cases in the good prognosis groups are reflected by a recurrence risk greater than zero in those groups. Test sensitivity and specificity were not explicitly included as parameters in the model because test performance was based on empirical analyses of recurrence risk for the good and poor prognosis groups for GEP and the NIH guidelines.

We assumed all women identified as "poor prognosis" received chemotherapy, all women identified as "good prognosis" did not receive chemotherapy, and all women lived at least 6 months after diagnosis. The long-term outcomes of these patients were then projected over their remaining lifetime using a Markov model (Fig. 2). A Markov model consists of mutually exclusive clinical outcome states (depicted as ovals), among which a patient can move each time period, in this case defined as 1 year.¹⁶ The possible clinical outcome states included in our model were: "no evidence of disease," "distant recurrence," and "death" (Fig. 2). Whether or not a woman has received chemotherapy as part of her treatment following diagnosis per GEP or NIH guidelines, all women began the time period 6 months after diagnosis responding to treatment ("no evidence of disease"). From "no evidence of disease," she could continue to respond to treatment, experience distant recurrence, or die of a non-breast cancer-related cause. Distant recurrence was considered to be progressive disease; that is, once a woman experienced distant recurrence, she could no longer transition to the "no evidence of disease" state. Additionally, we assumed that once a woman experienced distant recurrence, costs, and mortality risk were the same for all women regardless of whether they originally received adjuvant chemotherapy for the primary tumor.

We adopted a societal perspective for this analysis and followed the recommendations of the Panel on Cost-Effectiveness in Health and Medicine.¹⁷ We programmed the model in



Fig. 1. Diagram depicts decision tree portion of decision model for the time period from diagnosis to 6 months subsequent.



Fig. 2. Diagram depicts Markov model portion of decision model for the time period from 6 months after diagnosis until death, which shows all possible clinical outcomes and transitions between them.

a Microsoft Excel spreadsheet and validated it using decision analysis software (TreeAge. Release 4.0. Williamstown, MA: TreeAge Software, Inc., 2004). Study procedures were approved by the University of Washington Institutional Review Board.

Model parameters

The model parameters included probabilities of clinical events, costs, and utilities. Utilities are used to adjust survival for quality of life, and range from 0 (death) to 1 (perfect health).¹⁷ The value for the main analysis, range of values for the sensitivity analyses, and data sources for each parameter are shown in Table 1.

Table 1 Parameters used in decision model				
Parameter	Reference-case value (range)	Reference		
Probabilities (annual)				
Distant recurrence ^a				
Gene Expression Profiling Low Risk				
Years 1–10	0.018	9		
Years 11+	0.009	9		
Gene Expression Profiling High Risk				
Years 1–10	0.069	9		
Years 11+	0.007	9		
NIH Guidelines Low Risk				
Years 1–10	0.020	9		
Years 11+	0.000	9		
NIH Guidelines High Risk				
Years 1–10	0.047	9		
Years 11+	0.008	9		
Relative risk for distant recurrence from adjuvant chemotherapy	65% (±25%)	2		
Mortality due to distant recurrence	27% (±25%)	9, 52, 53		
Mortality from other causes ^b	0.01–0.62	18		
Utilities				
Post-diagnosis: no adjuvant chemotherapy	0.8 (0.6–1.0)	54		
Post-diagnosis adjuvant chemotherapy	0.5 (0.3–0.9)	54–56		
No evidence of disease	0.9 (0.8–1.0)	26, 54, 57–59		
Distant recurrence	0.3 (0.2–0.5)	26, 58–60		
Costs				
Gene expression profiling	\$3,460 (\$1,960–\$4,860)	20		
Adjuvant chemotherapy	\$21,984 (\$17,930-\$26,037)	Attributable cost analysis using reimbursement rates from a managed care organization, 33, 55, 61–63		
Distant recurrence	\$50,300 (\$40,300-\$60,300)	55, 56, 60, 61, 64–66		

^aRecurrence risk depended on prognosis group and duration of "no evidence of disease" state. ^bBackground mortality was estimated for 5-year age increments.

Probabilities of clinical events

We estimated the probabilities of clinical events from empirical analyses performed on patient-level data from the Netherlands Cancer Institute cohort, which was also the population for a validation study of GEP.9 We estimated the probabilities of a patient being identified as good prognosis or poor prognosis, the baseline risk of distant recurrence for each prognosis group (positive predictive value for poor prognosis groups, 1-negative predictive value for good prognosis groups), and the mortality associated with distant recurrence from Kaplan-Meier survival analyses. In the analysis of risk of distant recurrence, distant metastases as a first event was defined as a failure and other patients who experienced locoregional recurrence, a second primary cancer (including contralateral breast cancer) or death from causes other than breast cancer were censored. The Netherlands Cancer Institute cohort was a consecutive, population-based group of 295 breast cancer patients \leq 52 years of age. The median duration of follow-up was 6.7 years. The cohort had approximately equal proportions of women with lymph node negative and positive disease in the GEP and NIH good and poor prognosis groups, and the majority of the cohort (63%) did not receive adjuvant chemotherapy. Women in the GEP and NIH guidelines poor prognosis groups who received chemotherapy appeared to have more favorable survival compared to women who did not, which could have produced bias in the estimates of recurrence risk for GEP and the NIH guidelines. However, because there were approximately equal proportions of women who received chemotherapy in the GEP and the NIH guidelines poor prognosis groups, the recurrence risk estimates in the GEP and NIH guidelines poor prognosis groups should be equivalently biased, and should not affect the comparative recurrence risk estimates. The probability of death from causes other than breast cancer was obtained from US life tables.18

Risk reduction from adjuvant chemotherapy

Results from the Early Breast Cancer Trialists' Collaborative Group meta-analysis of adjuvant systemic therapy² suggest that for women < 50 years old with early stage breast cancer, chemotherapy reduces the 10-year risk of distant recurrence by 35%. We applied this risk reduction to the yearly recurrence risks for the poor prognosis groups (positive predictive value), which ensures that the chemotherapy benefit in reducing recurrence risk is applied only to those women in whom the disease would recur in the absence of chemotherapy. The way we incorporated the risk reduction from adjuvant chemotherapy in the model arises from the simple assumption that administering chemotherapy to a woman in whom the disease will not recur cannot reduce her risk of recurrence. Because GEP was designed and validated to identify women at high risk of distant recurrence, and because the prognosis of women with distant as compared to locoregional recurrence is less favorable,19 we did not consider the risk of local or regional recurrence. In the Netherlands Cancer Institute cohort, 4% of the women had locoregional recurrence, versus 34% with distant recurrence.

Economic costs

We included direct medical and nonmedical (time and transportation) costs in our calculations. A different GEP assay, targeted to lymph node–negative, ER+ women recently has been marketed for clinical use in the U.S. at a price of \$3460.²⁰ We used this price as an estimate for the cost of the GEP assay studied in our analysis, and varied the cost in sensitivity analyses. We assumed that it was costless to use NIH clinical guidelines, as tumor size and lymph node status, which form the basis for the guidelines, are routinely collected at pathology review.

Direct medical costs for breast cancer care were based on estimates from the published literature and empirical analyses. We assigned costs to distant recurrence and adjuvant chemotherapy and did not assign a cost for "no evidence of disease." To assign costs to distant recurrence, we incorporated point estimates for the annual/episodic costs of breast cancer treatment from the literature. We assigned a one-time cost to distant recurrence that represents the total cost from onset of distant recurrence until death. We considered studies that were (1) conducted in a U.S. health care setting, (2) estimated costs for the relevant clinical outcomes, (3) were published after 1990, and (4) explicitly described their methods. We found a limited number of studies of adjuvant chemotherapy costs in breast cancer; most costs were estimated for older chemotherapy regimens and there was substantial variation in reported results (\$5,000-\$16,000). We thus undertook an attributable cost analysis using managed care reimbursement data,²¹ and statistical methods that account for censored data.^{22,23} This approach estimated chemotherapy attributable costs as the difference between the cancer attributable costs of women with chemotherapy and women without chemotherapy. Briefly, we identified cases from a linked database of claims records from a health plan covering persons under age 65. Women with breast cancer in this plan were identified using the Surveillance, Epidemiology, and End Results (SEER) registry. Resource prices were based on reimbursements from the managed care organization. Women with chemotherapy in the database received a variety of agents and may have also received supportive care agents and/or hospitalization for complications from the chemotherapy. Because we computed incremental costs, the costs of adjuvant chemotherapy and distant recurrence were the only breast cancer treatment costs included in the model. Thus, we did not evaluate other treatment costs for the primary tumor, namely surgery and radiation therapy. All costs are represented in 2003 US dollars.

We estimated direct non medical costs by multiplying the estimated number of hours spent traveling to treatment facilities and in treatment (personal communication, Jeannine S. McCune, June 6, 2003) by the average hourly wage for women (\$13.48/hour) in the U.S.²⁴ Transportation costs were estimated based on typical miles traveled per visit (20 miles), cost per mile (36 cents, per IRS automobile mileage reimbursement rates), and cost of parking (\$2). The nominal cost of lost wages and transportation (\$300) was overwhelmed in the analysis by

Oestreicher et al.

the considerable costs of GEP and chemotherapy, so that a more accurate estimate of these costs was not necessary.

Quality of life (Utilities)

Utility weights were based on estimates reported in the literature. Utilities refer to the preferences individuals or society may have for any particular set of health outcomes and, in our study, allows for the adjustment of survival for quality of life.25 A utility may range from 0.0 (death) to 1.0 (perfect health). We considered those studies that were published after 1990 and used the standard gamble or time tradeoff methods to estimate utilities.17 We included studies that estimated utilities for clinical outcomes in the model, with explicitly described methods and results that were stated numerically. While it would have been preferable to utilize estimates obtained from communitybased samples,17 few such studies were available.26 We assigned utilities to the following clinical outcomes: 6 months after diagnosis, no chemotherapy; 6 months after diagnosis, with chemotherapy; "no evidence of disease" and "distant recurrence." The utility value for "6 months after diagnosis, with chemotherapy" is lower than the utility value for "6 months after diagnosis, no chemotherapy" because of the detrimental effects of potential toxicity from chemotherapy on quality of life. Additionally, although the utility value assigned to distant recurrence is quite low (0.3), we believe it is an appropriate value for a progressive disease state and is internally consistent with other utility values used in the model. Furthermore, utility for distant recurrence is not an influential parameter in our model, so potential bias in this parameter would not have a substantial impact on results.

Data analysis

Costs and quality-adjusted life years (QALYs) for future years were discounted at 3% per year. The incidence of distant recurrence, the incidence of breast cancer death, total QALYs, and direct medical and nonmedical costs were calculated for GEP and the NIH guidelines. To validate the model, the overall and recurrence-free survival outcomes derived from the model were compared with results from the Netherlands Cancer Institute cohort. Finally, the incremental cost-utility ratio, which is interpreted as the additional cost to provide one additional QALY, was calculated where appropriate. The numerator of the ratio is the difference in total costs between GEP and the NIH guidelines and the denominator is the difference in QALYs.

Sensitivity analyses and alternative testing strategies

We conducted one-way sensitivity analyses to evaluate the effect of varying individual probabilities, costs, and utilities on model results, while holding the others fixed. All parameters in the model were included in the sensitivity analyses (Table 1). In addition, multiway, probabilistic sensitivity analyses were performed using Monte Carlo simulation²⁷ and @Risk software (Palisade Corporation, Newfield, NY). For each simulation, the probabilities, costs, and utilities were randomly drawn from probability distributions that represented the uncertainty

of each of the model parameters and a cost-utility ratio was calculated. Ten thousand simulations were conducted to ensure convergence of results, thus removing "first-order" uncertainty, or within-individual predictive uncertainty.²⁸ We used logistic normal distributions for probabilities and utilities²⁷ and lognormal distributions for costs.²⁸ A mean costutility ratio and an "uncertainty interval" containing 95% of the values from the simulation were calculated.²⁸ The uncertainty interval provides an estimate of the overall uncertainty in the model due to uncertainty in all of the model parameters, or "second-order" uncertainty.²⁸

Because baseline recurrence risks in the GEP good and poor prognosis groups were not independent, we did not vary these parameters but instead evaluated the range of results that could be achieved using the current (Netherlands Cancer Institute) assay by changing the test cutoff to identify a tumor as poor prognosis.9 Changing the test cutoff alters the recurrence risk for each prognosis group. Finally, in addition to evaluating the testing of all women with either GEP or the NIH guidelines, we also considered the following alternative testing strategies entailing combined use of the results from GEP and NIH guidelines, or not using the results from NIH guidelines: (1) GEP for women identified as poor prognosis using NIH guidelines followed by chemotherapy for women who are identified as poor prognosis on both the NIH guidelines and GEP; (2) GEP for women identified as good prognosis using NIH guidelines followed by chemotherapy for women identified as poor prognosis on either the NIH guidelines or GEP; and (3) identifying 100% of women as candidates for adjuvant chemotherapy in lieu of using results from the NIH guidelines.

RESULTS

The NIH guidelines identified 96% of the cohort as high risk and thus candidates for chemotherapy, whereas GEP identified 61% of patients as high risk (Table 2). This prognostic categorization yielded sensitivities of 98% for the NIH guidelines and 84% for GEP. Specificities were 51% for GEP and 5% for the NIH guidelines. Accounting for the 35% risk reduction in dis-

Table 2			
Performance, costs, and outcomes of gene expression profiling vs. NIH			
guidelines			

8				
	Gene expression profiling	NIH guidelines	Difference	
Sensitivity	84%	98%	-14%	
Specificity	51%	5%	46%	
Proportion of women treated with chemotherapy	61%	96%	-35%	
Expected proportion of distant recurrences prevented	29%	34%	-5%	
Costs	\$29,754	\$32,636	-\$2,882	
Quality-adjusted life years	9.86	10.08	-0.21	

Genetics IN Medicine

tant recurrence resulting from chemotherapy, utilization of the NIH guidelines to identify and treat high-risk women with chemotherapy prevented 34% of women from experiencing distant recurrence compared to 29% for GEP. When the negative impact on life expectancy and quality of life from chemotherapy and distant recurrence were included, the NIH guidelines and GEP yielded 10.08 versus 9.86 QALYs, respectively. Total costs were \$32,636 for the NIH guidelines and \$29,754 for GEP. Because GEP produced lower QALYs and lower costs than the NIH guidelines, calculation of an incremental costutility ratio was not conducted.¹⁷

We validated the model by comparing the overall and recurrence-free survival outcomes and compared them to results from the Netherlands Cancer Institute cohort. We found that the model predicted approximately 7% fewer distant recurrence events than the women in the Netherlands Cancer Institute cohort. This difference is a result of accounting for censoring in the primary data by utilizing Kaplan-Meier survival analyses to estimate risks of distant recurrence and breast cancer mortality. If censoring, or incomplete follow-up time, were not accounted for, risks of distant recurrence would be overestimated.

We also considered the following alternative testing strategies entailing the combined use of results from the NIH guidelines and GEP or not using results from the NIH guidelines: (1) GEP as a confirmatory test for high-risk women identified by NIH guidelines; (2) GEP as a test to identify high-risk women missed by NIH guidelines; and (3) identifying 100% of women as candidates for adjuvant chemotherapy in lieu of using results from the NIH guidelines. Because the NIH guidelines had high sensitivity (96%), and the QALYs lost from failing to identify and treat women who would experience distant recurrence in the absence of chemotherapy was a strong driver of outcomes, the combined testing strategies were dominated by use of the NIH guidelines alone (data not shown) and were not considered further. Comparison of GEP versus identification of 100% of the Netherlands Cancer Institute cohort as candidates of adjuvant chemotherapy did not substantially increase QALYs but increased costs compared to results from the main analysis of GEP versus NIH guidelines (GEP vs. identification of all women: difference in QALYs = -0.17, difference in costs = -\$3942).

Results of sensitivity analyses

The tornado diagrams (Figs. 3 and 4) display the results of the one-way sensitivity analyses for the most influential parameters. QALYs were most sensitive to the test cutoff to identify a tumor as poor prognosis. In order for the GEP test to produce equivalent QALYs to the NIH guidelines, GEP sensitivity would need to be 95% or greater, while maintaining specificity (51%). Regardless of the test cutoff used to identify a poor prognosis tumor, the GEP assay studied in our analysis, at its current level of performance, does not attain a sensitivity of this magnitude. Total cost was most influenced by the costs of GEP and chemotherapy, and GEP was cost-saving over the entire range of variation of these parameters. The multi-way probabilistic sensitivity analysis demonstrated a 95% uncertainty interval of -0.32 to -0.09 for the difference in QALYs and -\$4686 to -\$1042 for the difference in costs.

DISCUSSION

We evaluated the potential clinical, patient, and economic benefits of a GEP assay versus the NIH guidelines to identify premenopausal women with early stage breast cancer for adjuvant chemotherapy. GEP identified 35% fewer women for chemotherapy than NIH guidelines, but the resultant quality of life benefits are outweighed by the decrease in life expectancy due to GEP's lower sensitivity. GEP's specificity was 10-fold higher than NIH guidelines, leading to lower overall costs. If GEP's sensitivity were to increase to at least 95% and its specificity (51%) was maintained, it would improve quality of life by allowing some women to safely avoid chemotherapy while at the same time not missing women whose survival is compromised by avoiding therapy. Additionally, our results suggest that although the cost of GEP is a major expense compared to the use of NIH guidelines, it appears that using the NIH guidelines may incur more chemotherapy costs, overwhelming the test costs of GEP and leading to the NIH guidelines being more costly overall compared to GEP. Furthermore, although the NIH guidelines identified the preponderance of the cohort (96%) as high-risk, our examination of alternative testing strategies suggests that identifying the entire cohort as high-risk in lieu of using the NIH guidelines would cost more and produce roughly equivalent QALYs compared to the main analysis of GEP versus the NIH guidelines.

There are several limitations to our analysis. First, in addition to the NIH guidelines, there are other guidelines applied in the United States that could have been utilized for the comparison, such as the National Comprehensive Cancer Network criteria.29 The National Comprehensive Cancer Network criteria are more conservative than NIH guidelines, so if we had compared these guidelines to the GEP assay instead of the NIH guidelines, these guidelines would tend to yield higher costs and have higher sensitivity than the NIH guidelines. Second, there is substantial uncertainty in several of our model parameters, most notably in the performance of the GEP assay. Although approximately equal proportions of women in the GEP and NIH guidelines poor prognosis groups received adjuvant chemotherapy, the type and duration of chemotherapy was not specified, which may have introduced bias in the estimates of recurrence risk. For example, if women in the GEP poor prognosis group received more intensive chemotherapy than women in the NIH guidelines poor prognosis group, it is possible that their recurrence risk was underestimated to a greater extent than for women in the NIH guidelines poor prognosis group, enhancing GEP's performance. In the good prognosis groups, a greater proportion of women received chemotherapy in the GEP group versus the NIH guidelines group (in fact no women had chemotherapy in the NIH guidelines good prognosis group), which would produce a bias in favor of GEP. However, because we found that GEP produces lower QALYs

Oestreicher et al.



Fig. 3. Diagram depicts the most influential model parameters in determining QALYs from most to least influential, and the effect of varying these parameters on total QALYs

than NIH guidelines, the presence of these potential biases would not change decisions about GEP based on our analysis. Additionally, although 77% of patients in the Netherlands Cancer Institute cohort were ER+, only 18% of these patients received adjuvant hormonal therapy (e.g., tamoxifen). Because adjuvant hormonal therapy is the current standard of care for women with ER+ tumors,5 if the cost-utility model were applied to current practice conditions, adjuvant hormonal therapy would reduce the baseline risk of distant recurrence in ER+ women and thus would decrease the absolute incremental benefit of adjuvant chemotherapy according to the proportion of ER+ patients in each prognosis group for GEP and the NIH guidelines. Because information on individual ER status for the GEP and NIH guidelines group was not available, we could not ascertain the direction of potential bias. Analogously, an increase in the efficacy of adjuvant chemotherapy for the GEP poor prognosis group would clearly favor the GEP strategy and there has been the suggestion that gene expression profiles may be able to predict response to chemotherapy.³⁰ Third, for our analysis we used one of several gene expression profiles that are marketed or in development for clinical use, and our results do not necessarily apply to these other tests. Like other profiles, the profile we used has not been extensively validated, but is in the process of further validation.

Fourth, we assumed all women in the poor prognosis groups received chemotherapy. In practice, some physicians may not follow the NIH guidelines in recommending chemotherapy or patients may refuse chemotherapy; thus, the use of chemotherapy may be lower than the values used in our analysis: 96% and 61% for the NIH guidelines and GEP, respectively.^{31,32} A similar decrease in compliance for GEP and the NIH guidelines would not likely have a dramatic effect on results. However, it is possible that compliance by patients and physicians may be different with GEP due to the perception of the value of genomic information: such issues merit further study.

Fifth, it is possible that the adjuvant chemotherapy cost we used in our model is an inflated estimate, as women who receive adjuvant chemotherapy are also more likely to receive radiation treatment.³³ To address uncertainty about the true cost of chemotherapy, we varied the cost of chemotherapy in the sensitivity analysis from \$17,930 to \$26,037, which would account for up to an 18% overestimate of the true cost of adjuvant chemotherapy. Based on previous report,³³ we believe the amount by which we varied the cost of adjuvant chemo-



Fig. 4. Diagram depicts the most influential model parameters in determining costs from most to least influential, and the effect of varying these parameters on total costs.

therapy in the sensitivity analysis should account for any potential bias in our chemotherapy cost estimate. Furthermore, although the cost of chemotherapy was an influential parameter in the model, the cost savings of GEP compared to NIH guidelines persisted over the entire range of uncertainty over which we varied chemotherapy cost.

Sixth, our cost-utility model may not fully capture all the elements that could influence physician-patient decisionmaking. Individuals may have different attitudes toward and perceptions of risk,34 which will affect their valuation of GEP and whether or not to undergo or recommend chemotherapy. The withholding of chemotherapy may cause undue anxiety for some patients, producing disutility for GEP as it relates to that patient. Health care providers may also be uneasy with a new risk assessment tool that results in fewer patients receiving chemotherapy. There is a burgeoning literature investigating how oncologists and patients address the benefits and risks of chemotherapy and the interpretation of prognostic information in breast cancer.35-46 Many of these studies suggest there is needed improvement in the communication of risks and benefits from provider to patient in cancer care. When genomic technologies are determined to have clinical utility in cancer care, physicians will need to know how to communicate information about risks and benefits in ways that promote informed patient decision-making. As GEP is developed for application in oncology and other clinical specialties, it will add to the genetics expertise required by practicing physicians, and thus to genetics educational needs.47-49

This analysis is an illustrative example of how genomic assays used in cancer care may be evaluated based on their accuracy and the downstream outcomes that result from it. Regardless of which GEP assay and clinical guidelines are used for comparison, how the natural history of breast cancer is modeled, and the values assigned to model parameters, it is evident that (1) test performance is crucial in determining outcome; and (2) the quality-adjusted life expectancy benefits of being spared from unwarranted chemotherapy are overwhelmed by the decreased life expectancy due to failing to identify and treat a woman who would experience distant recurrence. Although GEP has significant potential to provide clinical benefit, our study identifies current limitations in GEP test properties and suggests additional refinement and validation are needed before use in clinical practice.

As gene expression profiles are validated in well-designed clinical studies based on standardized protocols, their value as a prognostic factor may be enhanced. Currently, there are international trials underway for the use of this technology.^{50,51} Ultimately, the most informative studies may involve the evaluation of the incremental prognostic value of GEP over standard pathologic predictors, such as ER status, tumor grade, and proliferation markers such as ki-67 and mitotic count, routinely collected at the time of diagnosis. As additional gene expression profiles are developed, decision modeling will continue to provide a means to assess outcomes and identify key test properties.^{52–66}

ACKNOWLEDGMENTS

Nina Oestreicher was supported in part by the Institute for Public Health Genetics at the University of Washington and the Predoctoral Fellowship in Outcomes Research from the

Oestreicher et al.

PhRMA Foundation. These funding organizations had no role in the design and conduct of the study, the collection, management, analysis and interpretation of the data or the preparation, review or approval of the manuscript.

References

- Ries LAG, Eisner MP, Kosary CL et al, eds. SEER Cancer Statistics Review, 1975– 2001. Bethesda, MD: National Cancer Institute, 2004.
- Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 1998;352:930–942.
- Shapiro CL, Recht A. Late effects of adjuvant therapy for breast cancer. J Natl Cancer Inst Monog 1994;16:101–112.
- Shapiro CL, Recht A. Side effects of adjuvant treatment of breast cancer. N Engl J Med 2001;344:1997–2008.
- Eifel P, Axelson JA, Costa J et al. National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1–3, 2000. J Natl Cancer Ins 2001;93:979–989.
- National Comprehensive Cancer Network. Update: NCCN practice guidelines for the treatment of breast cancer. National Comprehensive Cancer Network. Oncology (Huntingt) 1999;13:41–66.
- Bast RC, Jr., Ravdin P, Hayes DF et al. 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. J Clin Oncol 2001;19:1865–1878.
- Sauter G, Simon R. Predictive molecular pathology. N Engl J Med 2002;347:1995– 1996.
- 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 2002;347:1999–2009.
- Paik S, Shak S, Tang G et al. A multigene assay to predict recurrence of tamoxifentreated, node-negative breast cancer. N Engl J Med 2004;351:2817–2826.
- King HC, Sinha AA. Gene expression profile analysis by DNA microarrays: promise and pitfalls. JAMA 2001;286:2280–2288.
- Agendia, 2004. Agendia announces a collaboration with MPI, Phoenix, AZ. Available at: http://www.agendia.com/common.asp?id=75&instantie=0&sContentID= 312-2. Accessed April 4, 2005.
- Ipsogen, 2005. Available at: http://www.ipsogen.com/products/index_breastcancer.html. Accessed April 4, 2005.
- Lyman G, Cosler L, Hornberger J. A 21-gene RT-PCR assay in lymph node negative, estrogen receptor positive early-stage breast cancer: an age-specific economic analysis. Paper presented at: San Antonio Breast Cancer Symposium. San Antonio, December 8–11, 2004.
- van 't Veer LJ, Dai H, van de Vijver MJ et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530–536.
- Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993;13:322–338.
- Gold ME, Russell LB, Siegel JE, Weinstein ME, eds. Cost-effectiveness in Health and Medicine. New York: Oxford University Press, 1996.
- National Center for Health Statistics. National Vital Statistics Report, Vol. 50, Number 15, 2002.
- National Cancer Institute, 2003. Cancer PDQ. Available at: http://www.nci.nih.gov/ cancertopics/pdq/treatment/breast/HealthProfessional/page8#Section_206. Accessed April 4, 2005.
- College of American Pathologists, 2005. Getting personal: Cancer assay could provide additional risk of recurrence. Available at: http://www.cap.org/apps/docs/cap_ _today/feature_stories/0204Gettingpersonal.html. Accessed April 4, 2005.
- Oestreicher N. The economic and clinical outcomes and policy implications of gene expression profiling in breast cancer care [dissertation]. Seattle: Department of Pharmacy, University of Washington, 2004.
- Hartunian NS, Smart CN, Thompson MS. The incidence and economic costs of cancer, motor vehicle injuries, coronary heart disease, and stroke: a comparative analysis. *Am J Public Health* 1980;70:1249–1260.
- Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. *Biometrics* 1997;53:419–434.
- Mishel L, Bernstein J, Boushey H. The State of Working America 2002/2003. Ithaca, NY: Cornell University Press, 2003.
- Drummond M, O'Brien B, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes. New York: Oxford University Press, 1999.
- Hall J, Gerard K, Salkeld G, Richardson J A cost utility analysis of mammography screening in Australia. Soc Sci Med 1992;34:993–1004.
- Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. *Med Decis Making* 1985;5:157–177.

- Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000;17:479–500.
- Carlson RW, Anderson BO, Bensinger W et al. NCCN Practice Guidelines for Breast Cancer. Oncology (Huntingt) 2000;14:33–49.
- Bast RC, Jr., Hortobagyi GN. Individualized care for patients with cancer: a work in progress. N Engl J Med 2004;351:2865–2867.
- Harlan LC, Abrams J, Warren JL, Clegg L, Stevens J, Ballard-Barbash R. Adjuvant therapy for breast cancer: practice patterns of community physicians. J Clin Oncol 2002;20:1809–1817.
- Du XL, Key CR, Osborne C, Mahnken JD, Goodwin JS. Discrepancy between consensus recommendations and actual community use of adjuvant chemotherapy in women with breast cancer. *Ann Intern Med* 2003;138:90–97.
- Barlow WE, Taplin SH, Yoshida CK, Buist DS, Seger D, Brown M. Cost comparison of mastectomy versus breast-conserving therapy for early-stage breast cancer. J Natl Cancer Inst 2001;93:447–455.
- Cher DJ, Miyamoto J, Lenert LA. Incorporating risk attitude into Markov-process decision models: importance for individual decision making. *Med Decis Making* 1997;17:340–350.
- 35. Glare P, Virik K, Jones M et al. A systematic review of physicians' survival predictions in terminally ill cancer patients. *BMJ* 2003;327:195.
- Baum M, Ravdin PM. Decision-making in early breast cancer: guidelines and decision tools. *Eur J Cancer* 2002;38:745–749.
- Simes RJ, Coates AS. Patient preferences for adjuvant chemotherapy of early breast cancer: how much benefit is needed? J Natl Cancer Inst Monogr 2001;30:146–152.
- Ravdin PM, Siminoff LA, Davis GJ et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol 2001;19:980–991.
- Loprinzi CL, Thome SD. Understanding the utility of adjuvant systemic therapy for primary breast cancer. J Clin Oncol 2001;19:972–979.
- Loprinzi CL, Ravdin PM, de Laurentiis M, Novotny P. Do American oncologists know how to use prognostic variables for patients with newly diagnosed primary breast cancer? J Clin Oncol 1994;12:1422–1426.
- Loprinzi CL, Laurie JA, Wieand HS et al. Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group. J Clin Oncol 1994;12:601–607.
- Ravdin PM. A practical view of prognostic factors for staging, adjuvant treatment planning, and as baseline studies for possible future therapy. *Hematol Oncol Clin North Am* 1994;8:197–211.
- McGuire WL, Tandon AK, Allred DC, Chamness GC, Ravdin PM, Clark GM. Prognosis and treatment decisions in patients with breast cancer without axillary node involvement. *Cancer* 1992;70:1775–1781.
- Siminoff LA, Fetting JH. Factors affecting treatment decisions for a life-threatening illness: the case of medical treatment of breast cancer. Soc Sci Med 1991;32:813–818.
- Fetting JH, Siminoff LA, Piantadosi S, Abeloff MD, Damron DJ, Sarsfield AM. Effect of patients' expectations of benefit with standard breast cancer adjuvant chemotherapy on participation in a randomized clinical trial: a clinical vignette study. J Clin Oncol 1990;8:1476–1482.
- Siminoff LA, Fetting JH. Effects of outcome framing on treatment decisions in the real world: impact of framing on adjuvant breast cancer decisions. *Med Decis Making* 1989;9:262–271.
- Burke W, Emery J. Genetics education for primary-care providers. Nat Rev Genet 2002;3:561–566.
- Burke W, Acheson L, Botkin J et al. Genetics in primary care: a USA faculty development initiative. *Community Genet* 2002;5:138–146.
- Burke W. Genetic testing in primary care. Annu Rev Genomics Hum Genet 2004;5: 1–14.
- Gianni L, Zambetti M, Clark K et al. Gene expression profiles of paraffin-embedded core biopsy tissue predict response to chemotherapy in patients with locally advanced breast cancer. Paper presented at: American Society of Clinical Oncology Annual Meeting. New Orleans, June 5–8, 2004.
- Rutgers EJ, Meijnen P, Bonnefoi H. Clinical trials update of the European Organization for Research and Treatment of Cancer Breast Cancer Group. *Breast Cancer Res* 2004;6:165–169.
- Giordano S, Buzdar A, Shu-Wan C, Hortobagyi G. Improvements in breast cancer survival: results from M.D. Anderson Cancer Center protocols from 1975–2000. Paper presented at: American Society of Clinical Oncology Annual Meeting. Chicago, May 18–20, 2002.
- Edwards BK, Howe HL, Ries LA et al. Annual report to the nation on the status of cancer, 1973–1999, featuring implications of age and aging on U.S. cancer burden. *Cancer* 2002;94:2766–2792.
- de Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van Der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. *Int J Cancer* 1991;49:538–544.

- Hillner BE, Smith TJ. Efficacy and cost effectiveness of adjuvant chemotherapy in women with node-negative breast cancer. A decision-analysis model. *N Engl J Med* 1991;324:160–168.
- Hillner BE, Smith TJ, Desch CE. Efficacy and cost-effectiveness of autologous bone marrow transplantation in metastatic breast cancer. Estimates using decision analysis while awaiting clinical trial results. JAMA 1992;267:2055–2061.
- Hayman JA, Fairclough DL, Harris JR, Weeks JC. Patient preferences concerning the trade-off between the risks and benefits of routine radiation therapy after conservative surgery for early-stage breast cancer. J Clin Oncol 1997;15:1252–1260.
- Hutton J, Brown R, Borowitz M, Abrams K, Rothman M, Shakespeare A. A new decision model for cost-utility comparisons of chemotherapy in recurrent metastatic breast cancer. *Pharmacoeconomics* 1996;9:8–22.
- Brown RE, Hutton J, Burrell A. Cost effectiveness of treatment options in advanced breast cancer in the UK. *Pharmacoeconomics* 2001;19:1091–1102.
- Cosler L, Hornberger J, Lyman G. Economic analysis of targeted chemotherapy using a 21-gene RT-PCR assay in lymph node negative, estrogen receptor positive

early-stage breast cancer. Paper presented at: American Society for Clinical Oncology Annual Meeting. New Orleans, June 5–8, 2004.

- Warren JL, Brown ML, Fay MP, Schussler N, Potosky AL, Riley GF. Costs of treatment for elderly women with early-stage breast cancer in fee-for-service settings. *J Clin Oncol* 2002;20:307–316.
- 62. Smith TJ, Hillner BE. The efficacy and cost-effectiveness of adjuvant therapy of early breast cancer in premenopausal women. *J Clin Oncol* 1993;11:771–776.
- Lokich JJ, Moore CL, Anderson NR. Comparison of costs for infusion versus bolus chemotherapy administration–Part two. Use of charges versus reimbursement for cost basis. *Cancer* 1996;78:300–303.
- 64. Taplin SH, Barlow W, Urban N, et al. Stage, age, comorbidity, and direct costs of colon, prostate, and breast cancer care. *J Natl Cancer Inst* 1995;87:417–426.
- Riley GF, Potosky AL, Lubitz JD, Kessler LG. Medicare payments from diagnosis to death for elderly cancer patients by stage at diagnosis. *Med Care* 1995;33:828–841.
- Fireman BH, Quesenberry CP, Somkin CP, et al. Cost of care for cancer in a health maintenance organization. *Health Care Financ Rev* 1997;18:51–76.