A Pilot Study to Establish a Clinical Model to Perform Phase II Studies of Breast Cancer Chemopreventive Agents in Women at High Risk with Biomarkers as Surrogate Endpoints for Activity

Vered Stearns,1,2 Ann Gallagher,1,2
Celina G. Kleer,6 Baljit Singh,2,3
Matthew Freedman,2 Bassem R. Haddad,2,4
Claudine Isaacs,1,2 Robert Warren,1,2
Martha Brown,7 Jennifer Cullen,5 Bruce Trock,2,5
and Daniel F. Hayes1,2

1Breast Cancer Program, Departments of 2Oncology, 3Pathology, 4Institute for Molecular and Human Genetics, and 5Cancer Prevention and Control, Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC, and Departments of 6Pathology and 7Internal Medicine, University of Michigan, Ann Arbor, Michigan

ABSTRACT

Purpose: Use of surrogate end point biomarkers in phase II trials may help select agents that appear to have activity and might be evaluated in future phase III definitive trials of breast cancer prevention. We performed a pilot clinical trial to establish the logistics for a clinical model to perform phase II studies of breast cancer chemopreventive agents in women at high risk with novel imaging techniques and candidate surrogate end point biomarkers for activity. We chose tamoxifen to establish proof of principal with a known effective agent.

Experimental Design: Women at a high risk of developing a new breast cancer and for whom tamoxifen was recommended were eligible. The women underwent baseline and 3 and 6 months mammogram and magnetic resonance imaging (MRI) of one breast to identify areas of water-like intensity (epithelial) and to determine the changes over time and MRI-directed core breast biopsies of these areas for surrogate end point biomarkers analysis.

Results: From August 1999 to March 2001, 26 women underwent baseline imaging and core biopsies. Sixteen women took tamoxifen and 10 chose not to. Overall, 79% of the samples contained glandular tissue evaluable for histology, but only 66% of the samples were evaluable for marker analysis. Only 12 patients had specimens with glandular tissue sufficient for marker analysis both at baseline and in at least one follow-up. Because of the small number of women with matched samples, marker analysis was not informative.

Conclusions: This study shows the feasibility of obtaining serial core breast biopsies from women at a high risk of developing a new breast cancer. Patient participation in this model is satisfactory, and such a model may provide indication of drug activity. MRI-directed biopsy did not provide a high yield of evaluable samples, and additional work on adequate collection of epithelial tissue for surrogate end point biomarker analysis is thus necessary.

INTRODUCTION

Identification of new measures to prevent breast cancer is critical. For many years, approaches to reduce breast cancer incidence included risk factor modification and bilateral prophylactic mastectomy. Unfortunately, many of the known breast cancer risk factors are not easily altered. Retrospective studies have shown that bilateral prophylactic mastectomy and/or oophorectomy are associated with a reduction in breast cancer incidence and mortality (1–3). However, these approaches are not acceptable to most women, and prospective randomized clinical trials will probably never be done. Given the limited and radical options, chemoprevention has become a desired intervention to reduce the risk of breast cancer.

The only agent currently approved for breast cancer risk reduction is the selective estrogen receptor (ER) modulator tamoxifen. Tamoxifen has been used for several decades in the treatment of metastatic breast cancer and as an adjuvant to primary therapy. Four prospective clinical trials have evaluated the role of tamoxifen as a chemopreventive agent (4–7). The largest trial reported to date is the Breast Cancer Prevention Trial conducted by the National Surgical Adjuvant Breast and Bowel Project P-1. In this trial 13,388 women were randomly assigned to 5 years of tamoxifen or placebo. With 52,401 person-years of follow-up, 2.6 and 1.3% of women in the placebo and tamoxifen arms, respectively, developed breast cancer (49% risk reduction, \( P < 0.00001 \); ref. 4). Tamoxifen was associated with a reduction of ER-positive tumors only.
Importantly, tamoxifen reduces the risk of future breast biopsy and benign breast diseases by ~30% (8). In the International Breast Cancer Intervention Study (IBIS-I), tamoxifen reduced the risk of breast cancer by 32% (7). Despite the large sample size and years of follow up, a survival advantage to women at high risk was not shown.

A second selective ER modulator, raloxifene, has been also studied as a potential chemopreventive agent. In the Multiple Outcomes of Raloxifene Evaluation, a significant reduction in breast cancer was observed in the raloxifene-treated group (relative risk = 0.35, 95% confidence interval = 0.21–0.58; ref. 9). As with tamoxifen, no reduction in ER-negative cancers was shown.

Approximately 20,000 women enrolled in the National Surgical Adjuvant Breast and Bowel Project P-2 trial (Study of Tamoxifen and Raloxifene) to evaluate reduction in breast cancer incidence, not mortality.

The retinoid, fenretinide, is the only non selective ER modulator agent tested against placebo in a phase III trial for prevention of a new breast cancer. In a prospective randomized clinical trial of 2972 women, fenretinide was not better than placebo in reducing incidence of new breast cancer (10). Thus, new agents for chemoprevention of breast cancer are needed, especially drugs that might inhibit ER-negative cancers. However, clinical end points, such as development of breast cancer or mortality, are infrequent in definitive clinical trials and may not occur for years. Definitive studies of new chemopreventive agents, for which clinical end points are considered necessary, require thousands or even tens of thousands of subjects and several years to complete.

Use of surrogate end point biomarkers in phase II trials may be helpful in selecting agents that appear to have activity and might be evaluated in future phase III definitive trials. Others have reported that random periareolar fine-needle aspiration may provide cytological diagnosis of hyperplasia or atypia and can also be used for surrogate end point biomarkers analysis (11). We initiated a pilot clinical trial designed to establish the logistics for a clinical model to perform phase II studies of breast cancer chemopreventive agents in women at high risk with novel imaging techniques to identify areas of the breast most likely to contain epithelial tissue and to target these areas for evaluation of candidate surrogate end point biomarkers for activity. In this pilot trial, we chose tamoxifen to establish proof of principal with a known effective agent.

PATIENTS AND METHODS

Patients. The prospective study was done within a formal protocol approved by the Georgetown University Medical Center’s Institutional Review Board. Eligible patients included women ≥35 years of age who were at an increased risk for a new breast cancer by meeting one or more of the following criteria: (a) risk of developing breast cancer in the ensuing 5 years was equal to or greater than that of a 60-year-old woman without other known risk factors (>1.6% over 5 years) as determined by the Gail model (12). If the patient had a strong family history of breast and/or ovarian cancer but her risk in accordance with the Gail model did not exceed 1.6% over 5 years, the Claus Risk Model was used to estimate risk (13); or (b) past history of biopsy-proven atypical ductal hyperplasia; past or present history of lobular carcinoma in situ; personal history of invasive breast cancer, treated with local therapy and for whom adjuvant therapy with tamoxifen only or no systemic therapy was appropriate; and previous or newly diagnosed ductal carcinoma in situ considering tamoxifen to reduce the risk of a new breast cancer. The women must have had at least one intact breast that had not been irradiated. Patients must have had adequate baseline hematologic, hepatic, renal, and coagulation function.

Patients who had previously taken tamoxifen for <2 years were eligible if they have been off tamoxifen for at least 1 year. Patients who had previously taken adjuvant tamoxifen for >2 years were not eligible. Patients may not have been taking hormone replacement therapy for at least 3 months before enrollment. Premenopausal women who chose to take tamoxifen were required to practice effective, nonhormonal, barrier contraceptive birth control throughout the study. Women with a current or previous cancer other than breast cancer, except nonmelanoma skin cancer, in the last 5 years were excluded unless they have been free of recurrent disease and off all anticancer treatment for at least 5 years. Other exclusion criteria included pregnancy or lactation, use of estrogen or other reproductive hormone replacement therapy, a prior history of thromboembolism, clotting disorders, a history of bleeding disorder, or anticoagulation use. Aspirin or nonsteroidal anti-inflammatory agent must have been discontinued for at least one week before biopsies.

Members of the Lombardi Comprehensive Cancer Center’s Breast Cancer Program informed the study coordinator of women who were potentially eligible for the study. The study coordinator also reviewed the schedules for breast-related radiology and surgery patients to identify potential participants. The study objectives and procedures were reviewed with each eligible patient during the clinic visit. All participants signed an Institutional Review Board-approved written informed consent.

Baseline and Follow-up Testing. Each patient underwent a baseline history and physical examination, including breast examination by one of the co-investigators. Baseline blood tests were done within 35 days before the MRI-directed biopsy and included a complete blood count, liver function tests, creatinine, prothrombin time, and partial thromboplastin time. Premenopausal women underwent a serum pregnancy test. Study blood samples were drawn for future studies of circulating markers of cancer risk or occult cancers. A history and physical, MRI-directed breast biopsy, and blood tests were repeated 3 and 6 months after enrollment. Women who received tamoxifen were also evaluated one month after starting the therapy. The 1-month evaluation included a brief history and physical, toxicity evaluation, a complete blood count, and liver function tests.

Breast imaging studies to determine epithelial density included a conventional mammography, consisting of the usual screening two-view assessment, and a breast MRI at baseline and 3 and 6 months after enrollment. Attempts have been made to schedule breast imaging and biopsy procedures 1 to 14 days after the last menstrual cycle for premenopausal women. Methods and results of breast imaging will be presented in a separate report. Initial methods for computer analysis of the breast MRI images have been reported previously (14).
Breast Biopsies. MRI was used to identify areas of increased water-like (epithelial) signal in the breasts of participating women. The area selected for biopsy was based on the region of the breast that showed delayed contrast enhancement. An initial diagnostic prone study was done to screen for occult malignancy and the patient was then repositioned supine for biopsy. The site chosen was estimated from a skin marker (colase capsule) placed near the likely target based on initial diagnostic sequence of MRI images and then re-imaged with the subject in position for biopsy. The study radiologist (M. Freedman) obtained core needle biopsies of these areas with a 14- to 18-gauge core needle biopsy apparatus at baseline and 3 and 6 months. Sampling of the normal breast tissue was often inconsistent with magnetic resonance compatible needles because the tissue was intermittently not completely severed by the needle. The problem was greatest with the 14-gauge needles, and therefore, different diameters and types of needles were used. The target number of samples obtained was 12, but additional samples were obtained if the tissue appeared visibly more likely to represent fat. Before the core biopsy, the breast was cleansed with alcohol and betadine and anesthetized with injection at the site chosen was estimated from a skin marker based on initial MRI images and then re-imaged with the subject in position for biopsy. The study radiologist (M. Freedman) obtained core needle biopsies of these areas with a 14- to 18-gauge core needle biopsy apparatus at baseline and 3 and 6 months. Sampling of the normal breast tissue was often inconsistent with magnetic resonance compatible needles because the tissue was intermittently not completely severed by the needle. The problem was greatest with the 14-gauge needles, and therefore, different diameters and types of needles were used. The target number of samples obtained was 12, but additional samples were obtained if the tissue appeared visibly more likely to represent fat. Before the core biopsy, the breast was cleansed with alcohol and betadine and anesthetized with injection at the target sites of one part bicarbonate and three parts 1% lidocaine with epinephrine. Several core samples were fixed in formalin-containing tissue fixation solution overnight and submitted to pathology for routine paraffin fixation. Two or three additional core samples were immediately placed in tissue freezing medium (OCT) in a cryomold and placed in dry ice with isopentane. Additional freezing medium was added until the cryomold was fully covered, and the block was then stored at −80°C. Finally, two or three core samples were placed in culture media and submitted for comparative genomic hybridization analysis to detect evidence of chromosomal aberrations. DNA used for comparative genomic hybridization was prepared either after a short-term culture of the core tissues or by direct extraction from the core samples as described below. Sections of the paraffin-embedded and frozen blocks were stained by H&E and evaluated by one of the study pathologists (B. Singh) for evidence of lesions that may require additional evaluation and/or therapy.

Therapy. A study co-investigator discussed with each eligible woman the advantages and possible adverse effects of tamoxifen therapy. Postmenopausal women were encouraged to participate in the National Surgical Adjuvant Breast and Bowel Project P-2 trial if they were eligible. Women who enrolled in P-2 were not eligible for this study. Women who were not eligible for P-2, those who were not willing to be randomized, and premenopausal women were encouraged to participate in this study regardless of their decision whether to take tamoxifen or not. Women, who chose not to take tamoxifen or wished to delay their decision, were eligible as long as they were willing to participate in the imaging and biopsy elements of the study. These women served as control subjects for comparison with those women who took tamoxifen. Tamoxifen was prescribed at 20 mg orally once daily.

Quality of Life. General health-related quality of life was assessed with the Medical Outcomes Study SF-12, which contains two subscales: the Mental Component Summary and the Physical Component Summary (15). The SF-12 scale is well validated and widely used among both medical patients and the general population (16). Each study participant completed the SF-12 survey at baseline and at the 3- and 6-month intervals. Each woman was also asked to complete the EuroQoL Linear Rating Scale, a self-rating standardized questionnaire of overall general health (17, 18). The EuroQoL-LRS asks a subject to rate their current overall health on a continuous scale from 0 (death) to 100 (best imaginable health). These measures were chosen for their ability to represent general health status in a basically healthy population, their responsiveness to the presence or absence of medical conditions, and their brevity and ease of completion.

Molecular Biomarkers. Paraffin-embedded tissues were sectioned into 4-μm thick sections and mounted onto separate slides. Markers that can be associated with response to therapy in general (apoptosis and proliferation) or those that may be modulated by tamoxifen specifically were evaluated (Table 1).

**Immunohistochemistry and Apoptosis Assays.** Slides were deparaffinized, hydrated, and subjected to antigen retrieval consisting of 10 mmol/L sodium citrate buffer (pH 6.0) at 95°C for 10 minutes. After endogenous peroxidase activity was blocked with peroxide block (BioGenex, San Ramon, CA), the slides were washed with PBS, and nonspecific antibody binding was blocked with normal goat serum (BioGenex) for 15 minutes at room temperature. Slides were washed with PBS and immunostained with commercially available antibodies (Table 1). Appropriate primary antibodies were diluted in primary antibody diluent (BioGenex) and applied to the slide. The slides were incubated at 37°C for 1 to 2 hours, washed in PBS, and...
incubated with secondary antibody (Link, BioGenex) for 20 minutes at room temperature. The slides were then washed in PBS and incubated with avidin biotin complex (Label, BioGenex) for 20 minutes at room temperature and washed with PBS. Slides were stained with 3,3'-diaminobenzidine and counterstained with hematoxylin and bluing solution (Optimax Wash Buffer, BioGenex). Finally, the slides were dehydrated and mounted. Immunohistochemistry was scored for percent positive cells and/or relative intensity by a study pathologist (C. Kleer) and a second investigator with light microscopy (Table 1).

Apoptotic index was determined with the terminal deoxynucleotidyl transferase-mediated nick end labeling method (TumorTACS In Situ Apoptosis Detection kit, Trevigen, Inc., Gaithersburg, MD). Slides were deparaffinized and hydrated, and endogenous peroxidase was removed. Slides were incubated in labeling buffer for 5 minutes and then for 1 hour at 37°C in a humid chamber with labeling solution containing terminal deoxyribonucleotidyl transferase, deoxyinosine triphosphate mix, and Mn2+. Slides were then transferred into stop buffer for 5 minutes and washed in PBS. Streptavidin-horseradish peroxidase was applied onto each sample for 10 minutes. Slides were washed in PBS twice, placed in 3,3'-diaminobenzidine for 20 minutes at room temperature and washed with PBS. Slides were stained with 3,3'-diaminobenzidine for 3 minutes, and counterstained with methyl green. Slides were dehydrated and mounted. Apoptotic index was calculated by counting and dividing the number of brown-stained nuclei by the total number of cells seen by light microscopy field at ×400 magnification by a study pathologist (C. Kleer) and a second investigator and are expressed as a percentage.

DNA Extraction and Comparative Genomic Hybridization Analysis. DNA used for comparative genomic hybridization was prepared either after a short-term culture of the core tissues or by direct extraction from the core samples. Short-term culture was done with routine protocols. Briefly, fresh core samples were collected in tissue culture medium and were immediately processed to establish a primary culture. Tissues were minced and treated with collagenase (0.125 mg/mL) for 1 to 2 hours. Cell cultures were set up in T25 flasks, and ~10 to 14 days later when ≥80% cell confluence was achieved, cells were trypsinized and used for DNA extraction with standard protocols. Core samples were frozen at −80°C immediately after the procedure until DNA extraction.

Comparative genomic hybridization was done as described previously (19). Briefly, nick translation was done to label tumor DNA and control DNA. Quantitative evaluation of hybridization was done with commercially available software (Applied Imaging, Pittsburgh, PA). Average ratio profiles were computed as the mean value of at least 8 ratio images and were used to identify changes in chromosome copy number.

Statistical Considerations. The objectives of the study were as follows: (a) to determine the success rate of accrual to a clinical model that included serial breast imaging and sampling for women at increased risk of breast cancer who were considering chemoprevention; (b) to determine the success rate of radiographic identification (by mammography and MRI) of areas of water-like intensity in the breasts of these women and to determine the changes over time of these areas; (c) to coordinate the logistics of using novel imaging techniques to direct core needle biopsies of high epithelial density in the breasts of these women to determine the success rate of obtaining epithelial tissue that can be evaluated for routine cytopathology and immunohistochemistry marker testing; (d) to evaluate baseline and change in markers in epithelial cells, including hyperplasia with and/or without atypia, apoptosis, proliferation, expression of ER, progesterone receptor, erbB-2, bcl-2, and p52; (e) to determine the effect of intervening short-term (6 months) treatment with tamoxifen on the appearance of epithelial high-density areas seen by special breast imaging and on epithelial biomarkers; and (f) to determine the success of using core samples to evaluate genomic changes with DNA prepared either directly from the cores or after a short-term culture of the cores. The DNA was analyzed with comparative genomic hybridization as a proof of principal.

Because of the pilot nature of the study, the statistics are principally descriptive. Ultimately, the changes in biomarker abnormalities before and after therapy were the critical end points, but sample sizes for these could not be estimated at the time of study design because of the lack of pilot study data. Therefore, we proposed to initially enroll 35 women in this study, of whom, we estimated that most will receive treatment with tamoxifen and a few will not.

To determine the success rate of accrual to the clinical model, the study coordinator documented the number of women who were asked to participate in this study. The number of women who agreed to participate in the study was divided by the total number invited to determine the percent that accepted enrollment.

To determine the success rate of obtaining epithelial tissue that can be evaluated for routine cytopathology and immunohistochemical marker testing (evaluable sample), we reported the number and percentage of core biopsies that contained sufficient material for evaluation. We also reported the number of women who had a baseline and at least one additional follow-up biopsy containing sufficient material for marker analysis (informative patient).

The difference between posttherapy and baseline values for a given marker were calculated. This was done for all control patients and for all tamoxifen-treated patients. Then, the mean of these differences was compared between the control and tamoxifen groups using the Student’s t test.

Descriptive statistics were computed for the Mental Component Summary, Physical Component Summary, and EuroQoL scales, including means, SDs, and ranges for all women, then stratified by tamoxifen use. Formal comparisons of Mental Component Summary, Physical Component Summary, and EuroQoL scores were made for women who chose to take tamoxifen versus not using the Student’s t test. The impact of outlying observation was addressed as necessary.

RESULTS

Accrual. From October 1999 to October 2001, 95 women were screened and 29 women registered to the study (31% enrollment). Reasons for nonparticipation included lack of agreement to undergo serial biopsies, unwillingness to discontinue hormone replacement therapy, or miscellaneous logistical reasons. Patient characteristics are summarized in Table 2. Median age was 48 years (range, 38–74 years). Over 50% of the
participants were premenopausal. Sixteen women (62%) chose to take tamoxifen whereas 10 did not. Half the study participants had a history of invasive carcinoma or ductal carcinoma in situ. Six women had lobular carcinoma in situ or atypical ductal hyperplasia. Other women were eligible based on Gail model criteria or family history.

**Collection of Breast Tissues.** Of the 29 patients who signed an informed consent, two declined to undergo the required study procedures and were excluded from analysis. An additional woman was found to be claustrophobic and underwent all study procedures, except MRI and MRI-directed biopsy and was not evaluable. Twenty-six women underwent baseline biopsy, and 23 (89%) and 22 (85%) women had 3- and 6-month biopsies respectively. Reasons for not undergoing biopsies are included in Table 3A.

Of the 26 baseline biopsies, 16 specimens (62%) contained normal glandular tissue, one specimen contained apocrine metaplasia, and one additional specimen contained usual ductal hyperplasia and fibrocystic changes. Eight samples did not contain glandular tissue. Overall, 56 of 71 biopsies contained material for histologic evaluation (79%), but a smaller portion (66%) contained sufficient material on serial sectioning for adequate marker analysis (Table 3B).

We defined informative patients as having a baseline biopsy and at least one follow-up evaluable biopsy specimen that contained sufficient epithelial cells for immunohistochemistry. Of the 26 participants, 7 women had three evaluable samples (baseline and 3 and 6 months), and 5 women had two evaluable samples (baseline and 3 or 6 months). Thus, the total number of informative patients was at least 12 (46%). Of these, eight took tamoxifen whereas four did not.

**Safety and Quality of Life.** No serious adverse events have been reported in the study. One patient complained of skin irritation at the biopsy site at the 6-month time point. Another patient noted a small scar at the baseline biopsy site. She was evaluated by the study radiologist who determined that an infection was not present and prescribed local measures. She returned for follow-up visits and biopsies. There were no reports of discomfort, hematomas, or other adverse events related to the study procedures.

Baseline overall mean Mental Component Summary, Physical Component Summary, and EuroQoL scores were 49.6, 53.7, and 87.3, respectively. Overall, no significant changes were observed after 3 and 6 months of study, suggesting that the study procedures did not adversely impact the patients. However, our sample size was too small to identify such changes (Table 4). Despite limited sample size, statistically significant differences were seen in EuroQoL-LRS at baseline and in Physical Component Summary score at month 3 such that women receiving tamoxifen had lower self-rated general health and lower physical health summary scores than those not on tamoxifen in all three respective time points. In this study, tamoxifen was not randomly assigned; rather, each participant made an individual decision whether she wished to take the drug or not. Thus, it is possible that baseline measures may differ among the groups. In addition, the women who took tamoxifen were also more likely to have already developed one breast cancer.

**Changes in Candidate Surrogate End Point Biomarkers.** We evaluated baseline and change in surrogate markers for women who received tamoxifen versus not. Unfortunately the number of informative women in each group was small. Despite the small sample size, there was a trend for reduction in progesterone receptor expression at the 6-month biopsy time point in tamoxifen-treated women; however, a change in ER was not observed. Fig. 1 shows baseline and change in ER and progesterone receptor expression in informative patients. Significant changes were not observed in HER2, proliferation or apoptosis. Other ER-regulated genes, including bcl2 and pS2, did not significantly change after short-term tamoxifen compared with no tamoxifen.

**Comparative Genomic Hybridization.** Forty-one samples were available for comparative genomic hybridization analysis. Thirty core samples were subjected to primary tissue culture. Of these, 25 cultures (83%) were successful, and comparative genomic hybridization analysis was successfully done in 23 cases. DNA was prepared from 11 frozen core samples, and comparative genomic hybridization analysis was successful in seven cases. No chromosomal alterations were detected in any of the 30 comparative genomic hybridization analyses.

**DISCUSSION.** We have conducted a prospective, pilot study designed to determine the willingness of women at high risk of a new breast cancer to participate in a study involving the use of potential surrogate markers of activity: radiographic evaluation of density changes and invasive collection of tissue for biomarker analysis. We conclude that women at a high risk for a new breast cancer are eager to participate in studies with serial imaging and tissue sampling. Indeed, for the most part, patient entry to our study was limited by research MRI time and not patient enthusiasm. Nearly one third of women who were screened agreed to participate, and of those who started

<table>
<thead>
<tr>
<th>Table 2 Patient characteristics (n = 29)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characteristics</strong></td>
</tr>
<tr>
<td>Age (y)</td>
</tr>
<tr>
<td>Median 48 (range, 38–74)</td>
</tr>
<tr>
<td>Menopausal status</td>
</tr>
<tr>
<td>Premenopausal</td>
</tr>
<tr>
<td>Perimenopausal</td>
</tr>
<tr>
<td>Postmenopausal</td>
</tr>
<tr>
<td>Risk</td>
</tr>
<tr>
<td>Invasive cancer</td>
</tr>
<tr>
<td>Ductal carcinoma in situ</td>
</tr>
<tr>
<td>Lobular carcinoma in situ</td>
</tr>
<tr>
<td>Atypical ductal hyperplasia</td>
</tr>
<tr>
<td>Gail model</td>
</tr>
<tr>
<td>Family history</td>
</tr>
<tr>
<td>Treatment group (patient choice)*</td>
</tr>
<tr>
<td>Tamoxifen</td>
</tr>
<tr>
<td>No tamoxifen</td>
</tr>
<tr>
<td>Baseline histopathology*</td>
</tr>
<tr>
<td>Normal glandular tissue</td>
</tr>
<tr>
<td>Usual duct hyperplasia, fibrocystic changes</td>
</tr>
<tr>
<td>Apocrine metaplasia</td>
</tr>
<tr>
<td>No glandular tissue</td>
</tr>
</tbody>
</table>

* Includes 26 patients who had baseline evaluation.
study procedures, only 1 of 26 (4%) withdrew subsequent consent to avoid future participation. This accrual fraction is substantially higher than the reported overall participation of patients with established cancer in clinical trials, which has been estimated at 3 to 5% (20). It is also superior to the rate of participation reported by us and others in a different chemoprevention clinical trial model directed toward women with established cancer (21, 22). The accrual is comparable with other trials that included women at a high risk of a new breast cancer with a similar design and are described below (23, 24).

It is feasible to obtain serial mammograms and MRI images, as well as serial core biopsies of water-like signal intensity areas on MRI. We showed that 89 and 85% of the women who underwent a baseline breast biopsy returned for a 3- and a 6-month biopsy, respectively. However, the yield of sufficient tissue evaluable for marker analysis was low. Although 79% of the biopsies revealed glandular tissue on H&E staining, many core biopsies did not contain sufficient glandular tissue (if any) for marker analysis (66% of biopsies evaluable). Needles for biopsy under magnetic resonance guidance are commercially available but are not ideal. Several different types and sizes of magnetic resonance compatible and not-compatible needles were tried. In general, we found that the magnetic resonance compatible needles were not as effective as the nonmagnetic resonance compatible needles in cutting the normal glandular tissue, particularly the 14-gauge size. Unlike tumor biopsies, the normal tissue samples often did not fill the needle well, resulting in the need for multiple sampling. The use of magnetic resonance to target high epithelial tissue is not recommended, and additional work on adequate collection of epithelial tissue is thus necessary.

### Table 4  Baseline and follow-up SF-12 and HRQoL scores

<table>
<thead>
<tr>
<th>Patients</th>
<th>Scale</th>
<th>Baseline (N = 17)</th>
<th>3 months (N = 17)</th>
<th>6 months (N = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>MCS</td>
<td>50 (10)</td>
<td>50 (12)</td>
<td>52 (9)</td>
</tr>
<tr>
<td></td>
<td>PCS</td>
<td>54 (8)</td>
<td>51 (10)</td>
<td>53 (7)</td>
</tr>
<tr>
<td></td>
<td>QOL</td>
<td>87 (10)</td>
<td>88 (9)</td>
<td>84 (23) *</td>
</tr>
<tr>
<td>No tamoxifen</td>
<td>MCS</td>
<td>51 (8)</td>
<td>48 (11)</td>
<td>54 (9)</td>
</tr>
<tr>
<td></td>
<td>PCS</td>
<td>57 (3)</td>
<td>56 (4)</td>
<td>54 (7)</td>
</tr>
<tr>
<td></td>
<td>QOL</td>
<td>94 (5)</td>
<td>90 (6)</td>
<td>93 (5)</td>
</tr>
<tr>
<td>Tamoxifen</td>
<td>MCS</td>
<td>48 (11)</td>
<td>51 (12)</td>
<td>51 (9)</td>
</tr>
<tr>
<td></td>
<td>PCS</td>
<td>51 (9)</td>
<td>47 (11)</td>
<td>53 (7)</td>
</tr>
<tr>
<td></td>
<td>QOL</td>
<td>84 (11)</td>
<td>87 (10)</td>
<td>79 (27) †</td>
</tr>
</tbody>
</table>

NOTE. At baseline and 3 months, 10 subjects reported tamoxifen use versus 7 not taking tamoxifen. At month 6, 11 subjects reported tamoxifen use whereas 7 were not taking tamoxifen.

Abbreviations: MCS, Mental Component Summary; PCS, Physical Component Summary; QOL, health-related quality of life (EuroQol-LRS).
* Includes a 2% outlier when excluded QOL 88 (11).
† When 2% outline excluded QOL 86 (14).
Other methods have been used to sample breast tissue or ductal fluid in women at high risk for a new breast cancer. The most successful method to date is random four-quadrant periareolar fine-needle aspiration. Random periareolar fine-needle aspiration allowed for cytological evaluation of hyperplasia or atypia that can predict the risk of developing a subsequent breast cancer (11). The method has also allowed for serial surrogate end point biomarkers determinations. However, a random periareolar fine-needle aspiration has a few shortcomings. For example, although it is considered "minimally invasive," random periareolar fine-needle aspiration requires a fairly rigorous sampling protocol, with 8 to 10 needle insertions in the breast and 4 to 10 needle passes within each needle insertion. Although performed under local anesthetic, this approach may be too invasive for many healthy women. Importantly, random periareolar fine-needle aspirations are successful when obtained by the same operator, which can be a challenge in most academic and non-academic centers. In addition, specimens obtained by random periareolar fine-needle aspiration are pooled so that they cannot be used to determine which part of the breast from which abnormal cells might have been collected.

Using directed fine-needle aspiration and/or core biopsies may allow to resample the same area from which the baseline samples were collected. Directed fine-needle aspiration or core biopsy should also minimize the risk to patients because it will decrease the number of needle insertions required when blind four-quadrant needle aspirations are done. Harper-Wynne et al. (25) have successfully used ultrasound to locate normal glandular breast to guide core-cut biopsies. Twenty-six of 29 patients had matched samples in the study; however, at least 2 patients required a second baseline core biopsy to obtain sufficient cellular material (i.e., 24 of 29 patients or 83% had matched evaluable biopsy before obtaining a second baseline biopsy; ref. 25). In addition, <50% of the cores obtained contained sufficient material, and in many cases, the authors had to add multiple specimens to identify a sufficient number of evaluable cells. In our study, up to 12 samples were removed at each biopsy; however, only a third was fixed in formalin and subsequently was paraffin-embedded, whereas other cores were frozen or subjected to short-term culture. It is possible that if we analyzed all 12 samples, we would have had a yield of evaluable biopsies similar to that reported by Harper-Wynne et al. (25). In an abstract reported by Baylor investigators, 63 of 265 eligible women (24%) with typical or atypical hyperplasia participated in a randomized study of tamoxifen versus placebo administered for 12 months. Forty-five paired specimens (71%) were available for comparison (24). However, data regarding the number of cores that have been removed are not available.

Others investigated ductal lavage to determine risk and to evaluate surrogate end point biomarkers. As with cytological evaluation of material obtained through random periareolar fine-needle aspiration, ductal lavage fluid may reveal atypia (26). However, the ability of ductal lavage to localize an area of abnormality, to provide histologic diagnosis, or to obtain sufficient material for multiple biomarker analyses has not been established. Ductal lavage material has been used to determine candidate gene methylation, a method that requires only a few cells on a single slide, an assay that may be used to predict subsequent risk (27).

Only a few phase II breast cancer chemoprevention studies that evaluated surrogate end point biomarkers before and after an intervention have been reported, and the experience with these has been suboptimal. To our knowledge, only a single study has been completed and reported in which a novel, putative chemopreventive agent was tested in a prospective controlled trial, including high-risk women without known breast pathology, other than atypia (23). In this study, 401 women were potentially eligible and 119 (30%) were randomized to 6 months of treatment with either difluoromethlornithine or placebo. Of those, 114 women (96%) completed the study during which they agreed to undergo random, periareolar fine-needle aspiration cytology. Unfortunately, no changes were detected in any of the surrogate markers evaluated in this trial. Nonetheless, as in our study, participation was brisk, and no major toxicities or poor outcomes were observed. Others evaluated the effects of hormonal treatments on breast biopsies obtained from women...
with benign breast diseases. In the study conducted by the Baylor investigators, preliminary results showed decrease in ER and progesterone receptor and increase in bcI2 in normal and hyperplastic tissue of tamoxifen-treated women (24). In the study of 32 women with a history of benign breast disease, ductal carcinoma in situ or lobular carcinoma in situ who underwent core breast biopsies before and after 3 months of the aromatase inhibitor letrozole, Harper-Wynne et al. (25) did not observe significant changes in Ki67 or ER.

We and others have done pilot studies in women with established breast epithelial in situ or invasive malignancies in which a putative chemopreventive agent is administered between diagnostic and definitive surgery. Accrual to these studies, at least in the United States, has been more difficult, perhaps because of the high level of anxiety facing such women at the time of diagnosis (21, 22). For example, only 37 of 267 invited women (14%) agreed to participate in such a trial investigating the effects of perillyl alcohol on breast tissue (22). In addition, whether an agent-induced modulation of surrogate end point biomarkers in malignant cells can predict the activity of the agent as a chemopreventive agent is not known.

A secondary objective of the study was to generate preliminary data regarding chromosomal defects in histologically normal tissue from high-risk women. However, no alterations in the DNA copy number were detected in the 30 evaluated samples, including those collected before or after tamoxifen therapy whether the DNA was extracted directly from a frozen core sample or a short-term culture. These negative results suggest that normal appearing epithelium from high-risk women may not yet have developed gross genetic defects. However, arguably, comparative genomic hybridization may not be sufficiently sensitive to detect finer genetic abnormalities, and our data may be confounded by the presence of DNA from nonepithelial cells in the assay. Several previous studies, including our own, suggest that comparative genomic hybridization is capable of detecting chromosomal aberrations when present in ~50% or more of the cells (28).

In conclusion, we found that women at high risk for breast cancer enthusiastically participated in this invasive phase II trial and that there were few if any adverse effects of their doing so. Ongoing trials are designed to improve and validate collection and analysis of surrogate end points, including breast density reduction and changes in ductal epithelial biomarker expression. Given the early but limited success of chemoprevention with selective ER modulators and the many potential chemopreventive agents that could be tested, additional evaluation of phase II clinical trial models with surrogate radiographic and biomarker indications of activity should lead to reliable identification of new drugs that could be tested in large-scale, definitive phase III studies.

ACKNOWLEDGMENTS

We thank Ahmad Farhad, Solomon Kebede, Anita Sarcone and Michele LeBlanc for technical support. We also thank Dr. Mark Rubin for advice with optimization of methods.

REFERENCES

tional Breast Cancer Intervention Study (IBIS-I): a randomised preven-
tional Breast Cancer Intervention Study (IBIS-I): a randomised preven-


A Pilot Study to Establish a Clinical Model to Perform Phase II Studies of Breast Cancer Chemopreventive Agents in Women at High Risk with Biomarkers as Surrogate Endpoints for Activity

Vered Stearns, Ann Gallagher, Celina G. Kleer, et al.


**Updated version** Access the most recent version of this article at: [http://clincancerres.aacrjournals.org/content/10/24/8332](http://clincancerres.aacrjournals.org/content/10/24/8332)

**Cited articles** This article cites 27 articles, 10 of which you can access for free at: [http://clincancerres.aacrjournals.org/content/10/24/8332.full.html#ref-list-1](http://clincancerres.aacrjournals.org/content/10/24/8332.full.html#ref-list-1)

**Citing articles** This article has been cited by 6 HighWire-hosted articles. Access the articles at: [http://clincancerres.aacrjournals.org/content/10/24/8332.full.html#related-urls](http://clincancerres.aacrjournals.org/content/10/24/8332.full.html#related-urls)

**E-mail alerts** Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions** To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.