AD_____

Award Number: W81XWH-04-1-0570

TITLE: Endocrine Therapy of Breast Cancer

PRINCIPAL INVESTIGATOR: Robert Clarke, Ph.D.

CONTRACTING ORGANIZATION: Georgetown University
Research and Technology Development Service
Washington, DC 20007

REPORT DATE: June 2008

TYPE OF REPORT: Annual

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. Agency Use Only (Leave blank)	2. Report Date	3. Report Type and Period Covered (i.e., annual 1 Jun 00 - 31 May 01)
	16 Jun 2008	Annual: 17 MAY 2007 - 16 MAY 2008

4. Title and Subtitle
Endocrine Therapy of Breast Cancer

5. Award Number
W81XWH-04-1-0570

6. Author(s)
Robert Clarke, Ph.D.

7. Performing Organization Name (Include Name, City, State, Zip Code and Email for Principal Investigator)
Georgetown University
Research and Technology Development Service
Washington, DC 20007

E-Mail: clarker@georgetown.edu

8. Performing Organization Report Number (Leave Blank)

9. Sponsoring/Monitoring Agency Name and Address
U.S. Army Medical Research and Materiel Command,
Fort Detrick, Maryland 21702-5012

10. Sponsoring/Monitoring Agency Report Number (Leave Blank)

11. Supplementary Notes (i.e., report contains color photos, report contains appendix in non-print form, etc.)

12a. Distribution/Availability Statement (check one)
XX Approved for public release; distribution unlimited

12b. Distribution Code
(Leave Blank)

13. Abstract *(Maximum 200 Words)* *(abstract should contain no proprietary or confidential information)*

A recent controversy in the treatment of estrogen receptor positive (ER+) breast cancers is whether an aromatase inhibitor, *e.g.*, letrozole (LET) or TAM should be given as first line endocrine therapy. Unfortunately, response rates are lower, and response durations are shorter, on crossover than when these agents are given as first line therapies, *e.g.*, ~40% of tumors show crossresistance to TAM or an aromatase inhibitor on crossover. Only 50% of ER+ tumors respond to endocrine therapy. Currently, we fail to predict endocrine responsiveness in about 66% of ER+/PgR- (progesterone receptor), 55% of ER-/PgR+, and 25% of ER+/PgR+ tumors. In this new Clinical Translational Research Award, we hypothesize that our analytical methods can extract expression profiles of breast tumors that define their responsiveness (sensitive *vs.* resistant) to endocrine therapy. These profiles, when combined with known predictive/prognostic factors, will support neural network and biostatistical classifiers or committee machines that predict each tumor's endocrine responsiveness. Our objectives are to array breast cancer cases, build classifiers of endocrine responsiveness (using microarray data), and validate these classifiers in independent data sets. In the long term, we will design custom arrays for use in clinical practice. Genes will be further studied using cellular and molecular methods, and their role as therapeutic targets explored.

14. Subject Terms (keywords previously assigned to proposal abstract or terms which apply to this award)
Antiestrogen, aromatase inhibitor, anastrazole, bioinformatics, biomarkers, biostatistics, breast cancer, class prediction, clinical trial, computer science, engineering, immunohistochemistry, letrozole, microarrays, molecular profiling, neural networks, recurrence, resistance, tamoxifen.

15. Number of Pages: 10

17. Security Classification of Report	18. Security Classification of this Page	19. Security Classification of Abstract	20. Limitation of Abstract
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

TABLE OF CONTENTS

1. Xuan, et al. *Open Appl Informatics,* 1: 11-19, 2007.
2. Ressom, et al. *Front Biosci,* 13: 691-708, 2007.
3. Gomez, et al. *FASEB J,* 21: 4013-4027, 2007
4. Naughton, et al. *Mol Endocrinol,* 21: 2615-2626, 2007.
5. Gong, et al. *Gene Regulat Sys Biol*, 1: 349-363, 2007.
6. Wang, et al. *BMC Bioinformatics* 9: 1-9, 2008.
7. Clarke, et al. *Nature Rev Cancer,* 8: 37-49, 2008.

INTRODUCTION

Endocrine therapy is often the least toxic and most effective treatment for hormone receptor positive invasive breast cancer. Such therapy includes antiestrogens (tamoxifen, fulvestrant) and aromatase inhibitors (anastrozole, letrozole, exemestane). Tamoxifen (TAM) increases disease free and overall survival in the adjuvant setting, reduces the incidence of estrogen receptor positive disease (ER+; unless otherwise noted ER=ERα) in high-risk women, and reduces the rate of bone loss secondary to osteoporosis in postmenopausal women [1,2]. Aromatase inhibitors are effective only in the absence of functioning ovaries - TAM can be used regardless of menopausal status. Recent studies suggest that anastrozole may be superior to TAM in the adjuvant treatment of postmenopausal women with ER+ breast cancer; other studies report higher overall response rates with letrozole (LET) *vs.* TAM as first line therapy in the metastatic setting. Thus, a recent controversy in the management of patients with ER+ disease is whether an aromatase inhibitor or TAM should be given as first line endocrine therapy [3-9].

In this Clinical Translational Research award, we will build classifiers that accurately separate antiestrogen sensitive from antiestrogen resistant breast tumors and begin to assist in the direction of specific endocrine treatments (antiestrogen *vs.* aromatase inhibitor) to *individual* patients. We hypothesize that endocrine responsiveness is affected by a gene network, rather than the activity of only one or two genes or signaling pathways [10-12]. Since the key components of such a network are unknown, we must study 10,000s of genes. We will use Affymetrix GeneChips. We will not identify mutational events, the presence of mRNA splice variants, or post-translational protein modifications. However, these factors have major effects on the transcriptome and their "footprints" should be identified by expression microarrays.

BODY

Overview: We will build classifiers that separate antiestrogen sensitive from antiestrogen resistant breast tumors and begin to assist in the direction of specific endocrine treatments (antiestrogen *vs.* aromatase inhibitor) to *individual* patients. To achieve this goal, and consistent with a CTR award, we will complete a 4-year, prospective, neoadjuvant study with Letrozole (LET) or TAM as the only systemic therapy. We will obtain molecular profiles from Affymetrix GeneChips and further develop and apply our innovative bioinformatic and biostatistic methods to explore these high dimensional data sets and build/validate new classifiers. A more accurate predictor of endocrine responsiveness would have widespread clinical use, allowing women and physicians to make more individualized and appropriate treatment decisions. For example, patients with tumors predicted to be resistant to antiestrogens and/or aromatase inhibitors would be strong candidates for an early intervention with cytotoxic chemotherapy.

In most predictive/prognostic marker studies investigators focus on a *single* factor and whether they obtain a p-value that reaches conventional statistical significance. Our approach is different because we will determine whether we can find joint gene subsets that can separate patients into sufficiently distinct groups that should differ in their treatment. We will **(1)** analyze >33,000 genes on retrospective and prospective material, **(2)** apply new biostatistical and bioinformatic methods to identify ~40 potentially informative "biomarkers," **(3)** build neural network and biostatistical model classifiers, **(4)** evaluate the joint discriminant power of selected genes concurrently rather than as single biomarkers, **(5)** focus on prediction for individual patients where the assessment of a p-value is less important than the classification rate of our predictors, **(6)** validate the classifiers in independent data sets, and **(7)** explore the ability of predictors to refine the targeting of *specific* endocrine therapies.

Evidence has begun to accumulate suggesting that an aromatase inhibitor might be a more effective first line endocrine therapy for some breast cancer patients than the current standard of care (Tamoxifen). These data have generated considerable interest and controversy, in part because unlike TAM, there are no long term studies with aromatase inhibitors where definitive survival data are available. Our study could provide new and

innovative insights into how to approach the more effective targeting of specific endocrine therapies to individual patients.

Specific Aims
We will complete two clinical studies and collect gene expression profiles from which to build predictors of endocrine responsiveness. Predictors will be built in Specific Aim 2 and validated in Specific Aim 3.

AIM 1: Clinical Studies - **Clinical Study-1** (retrospective) is of pretreatment, single, frozen samples where we will compare the molecular profiles of tumors that recurred on TAM with those of tumors that did not recur. Each resistant sample is matched with a TAM sensitive sample by age, stage, and duration of follow-up. We also have further, single (unmatched), frozen samples from patients already progressing on TAM. **Clinical Study-2** is a prospective study of breast tumor samples from patients treated with neoadjuvant TAM or LET.

AIM 2: We will develop and apply novel bioinformatics and biostatistics to discover gene subsets that define the molecular differences between endocrine sensitive and resistant breast tumors. These genes will be used, in combination with established predictive/prognostic factors, *e.g.,* ER, PgR, stage, to build innovative classifiers that can better predict an individual tumor's endocrine responsiveness.

AIM 3: We will test, optimize, and validate the performance of the classifiers from Aim 2 in retrospective studies of human breast tumors. We will measure each gene individually by IHC, *in situ* RNA hybridization (ISH), or *real time* PCR (RT-PCR).

KEY RESEARCH ACCOMPLISHMENTS

As noted in previous reports, progress on the clinical goals for this award was greatly delayed because of the time taken to obtain DOD approval of our preexisting institutionally approved IRBs at Georgetown University and at the University of Edinburgh. All institutionally approved protocols and requested material were submitted to the DOD in July 2004; additional information was requested by the DOD several months later and submitted in November 2004. We did not receive final approval to proceed with the clinical studies until March 2005. Much of this delay seems to have been entirely unavoidable (see prior reports). We continue to make significant strides in our development of new analytical procedures. Publications supported since the commencement of this award are listed under "Reportable Outcomes"; these constitute some of our major accomplishments in the past year. These and other key research accomplishments are presented below.

Progress on our Statement of Work

- **TASK 1.** Array breast tumor samples from Clinical Studies 1 (retrospective) and 2 (prospective)

We have received n=481 breast specimens from breast cancer patients treated with endocrine therapy (or not, *i.e.,* surgery and radiation only in selected retrospective cases) as described in the original application. These specimens represent a mix of the initial prospective and retrospective specimens. All of these specimens have now been fully analyzed and annotated by the study pathologist. We have successfully extracted total RNA from 480 specimens and labeled 300 for analysis. We have also completed the hybridization and assessment of microarray data quality control on over 200 breast cancer specimens.

We requested that the specimens be sent independent of the clinical information, so that we could adequately and appropriately randomize the RNA preparation, labeling and hybridization and minimize any operator-induced or technology-induced bias. All specimens were processed using our standard operating procedures; each manipulation being performed by the same individual to further reduced inter-operator variability. Details

of the methods, quality control measures and general experimental approaches have been described in detail in earlier annual reports.

We have also found these data to be particularly useful in supporting other studies that are ongoing in the laboratory. For example, these data have been used to support R01 applications on genes we identified and described in the preliminary data for this application. One of these has received a competitive score and will soon be revised and resubmitted. As in prior years, we have also used these data to provide preliminary data on gene expression values that have led to our colleagues initiating other studies directed at developing therapeutic strategies to target individual genes we have identified from within this data set or from other sources.

We continue to array specimens as we obtain the appropriate clinical information from Scotland. In this regard, we have tended to prioritize retrospective study material because these have definitive clinical outcomes (survival). As noted for Task 2 (below), we have begun analysis of the retrospective study and report below the results of our initial studies. For the prospective study, we continue to obtain outcomes data and array specimens as the clinical information dictates (we array specimens when the clinical information is sufficiently informative to be included in our analyses). Since many endocrine treated breast cancers tend to recur in later years, it is to be expected that we may obtain our most valuable data (recurrence) after this award has ended. However, analysis using clinical response as the outcome measure will proceed for the prospective clinical study as was proposed in the original application.

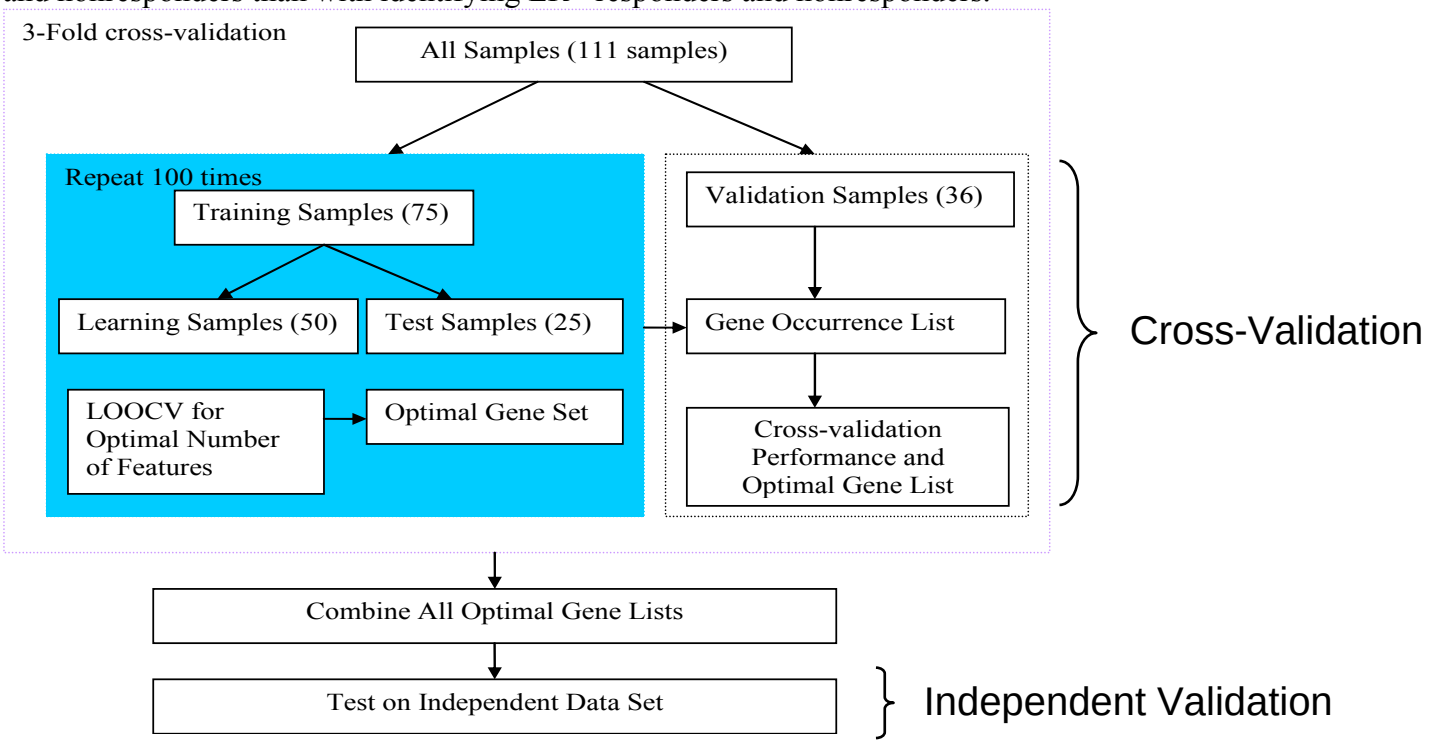
- **TASK 2.** Store, process, and train/optimize classifiers from gene expression microarray data (modified to reflect our adoption of caArray and other caBIG tools)

As noted in our previous reports, we also continue to make significant progress on addressing this task, largely as a consequence of our involvement in the National Cancer Institute Center for Bioinformatics (NCICB) led caBIG project. The PI (Dr. Clarke) leads the Lombardi Comprehensive Cancer Center's caBIG team and we have been actively involved in the development of caArray (NCICB's grid-enabled, MIAME compliant, microarray database). We successfully hosted a major caBIG face-to-face meeting between the caBIG Architecture and Vocabulary and Common Data Elements Workspaces at the Lombardi Comprehensive Cancer Center at Georgetown University in January 2008.

We also continue to further develop and optimize our data analysis algorithms, with particular success in the design of new approaches for network analysis. We have found approaching this goal to be realistic in a much shorter time-frame than initially expected and have already submitted several manuscripts for publication. We also continue to improve our existing algorithms for data analysis. Relevant publications in this area are included below in the section "Reportable Outcomes."

We have now acquired sufficient data for initial analysis of the endocrine therapies and outcomes. These were presented as an oral presentation at the recent "Era of Hope" meeting in Baltimore, Maryland. For this study, we used the data from our Edinburgh data set (BC030280 data set) to generate classifiers of clinical outcome and validated these classifiers using published data sets. Comparison of the recurrence data across a 25 year period show that our dataset, unlike others in ER+ breast cancers, has strong representation of the late recurrences (>10 years) characteristic of ER+ disease. Other existing data sets appear biased by a high proportion of early recurrences (<10 years) that may be more representative of poor prognosis cases. While early recurrences are also present in our dataset, we hope to have a sufficiently representative distribution to be able to compare early vs. late and possibly obtain a more definitive assessment of endocrine responsiveness. Moreover, our dataset is predominately ER+, allowing us specifically to address our central hypothesis of predicting responsiveness in these patients. Other data sets also tend to include large subsets of ER- cases,

which can bias the classifiers by identifying features associated more with separating ER- and ER+ responders and nonresponders than with identifying ER+ responders and nonresponders.



The data analysis design is shown above (LOOCV = leave-one-out cross validation) for our initial study. Independent validation was explored in three published data sets:
(i) Ma et al., PNAS 100: 5974, 2003 and Chanrion et al. Clin Cancer Res 14: 1744, 2008 have endocrine treated cases and are useful to assess potential predictive ability (responsiveness to endocrine therapy)
(ii) Wang et al. Lancet 365:671, 2005 is a dataset for assessing prognostic ability (outcome independent of treatment). All datasets used the same gene expression microarray platform we used in generating the BC030280 data set.

We built and tested our *initial* classifiers on the BC030280 data set. Performance was evaluated against a series of three preset benchmarks (≥70% performance in accuracy, sensitivity, specificity), requiring our classifiers to exceed at least two benchmarks in the two endocrine treated independent data sets (group (i) above). Accuracy is the percentage of cases called correctly (recurred; did not recur); sensitivity and specificity as estimated from the receiver operating characteristic (ROC) curve as described in he original application. Our initial classifiers built on the BC030280 data set met the benchmarks for the two endocrine data sets but failed on the prognostic data set. This is very encouraging as it suggests that our classifiers may be more accurate in predicting endocrine responsiveness than simply driven by poor prognosis. We are currently working to optimize these classifiers and including additional cases as they are arrayed. Thus, we believe that progress on Task 2 is fully consistent with our initial goals.

- **TASK 3.** Retrain/reoptimize classifiers using IHC data from Series 1 (Archival Tissues) and Series 2 (Scottish Adjuvant TAM Trial) for Validation

To perform this task we will obtain clinical information and breast tumor samples from University of Edinburgh (formalin fixed/paraffin embedded). We will rank and prioritize selected joint genes from RNA classifier built and optimized in TASK 2 (above) and retrain/reoptimize the initial neural network IHC classifier (MLP). Finally, we will validate IHC classifier on independent data sets (data sets not used to build and train the MLP classifiers).

As acknowledged in prior reports, we remain unable to move this task substantially forward on the timeframe as initially proposed because of the delays in getting approval to work with the clinical specimens (Task 3 cannot begin until Tasks 1 and 2 are almost complete). Nonetheless, we will work towards addressing this aim where possible in the next year.

REPORTABLE OUTCOMES

Papers and Meeting Reports*

New Publications (for the present reporting period)

- Clarke, R., Ressom, H., Wang, A., Xuan, J., Liu, M.C., Gehan, E. & Wang, Y. "The properties of very high dimensional data spaces: implications for exploring gene and protein expression data." *Nature Rev Cancer,* 8: 37-49, 2008.

- Gomez, B.P. Riggins, R.B., Shajahan A., Klimach, U., Zhu, Y., Zwart, A., Wang, M., Wang, A. **&** Clarke, R. "Human X-box binding protein-1 confers both estrogen-independence and antiestrogen resistance in breast cancer." *FASEB J*, 21:4013-27, 2007.

- Xuan, J., Wang, Y., Clarke, R. & Hoffman, E., "An iterative nonlinear regression method for microarray data normalization," *Open Appl Informatics J*, 1: 11-19, 2007.

- Ressom, H.W., Varghese, R.S., Zhang, Z., Xuan, J. & Clarke, R. "Classification algorithms for phenotype prediction in genomics and proteomics." *Front Biosci*, 13: 691-708, 2008

- Naughton, C., Kuske, B., MacLeod, K., Clarke, R. Cameron, D.A. & Langdon, S.P. "Progressive loss of estrogen receptor alpha (ERα) cofactor recruitment in endocrine resistance." *Mol Endocrinol,* 21:2615-26, 2007

- Gong, T., Xuan, J., Wang, C., Li, H., Hoffman, E., Clarke, R. & Wang, Y. "Gene module identification from microarray data using nonnegative independent component analysis." *Gene Regulat Sys Biol,* 1: 349-363, 2007.

- Wang, C., Chen, L., Zhao, P., Hoffman, E., Clarke, R., Wang, Y. & Xuan, J. "Motif-directed component analysis for regulatory network inference." *BMC Bioinformatics,* 9: S21, 1-9, 2008.

*We include in the appendix reprints of those papers that are already published and for which we have proofs or reprints. We do not list here or include in the appendices any published abstracts, but can do so if requested. Several other manuscripts also are submitted and in preparation – these will be cited reported in the next report at the end of our no cost extension.

Comment on Subcontracts: Please also note that the majority of our publications here and in prior years include coauthors from one or both of our subcontracts. Thus, our program is working very effectively and collaboratively, this should further be apparent in the development of new informatics methods (Virginia Polytechnic and State University subcontract) and the large number of high quality breast tumor specimens we have obtained from the University of Edinburgh.

CONCLUSIONS

We continue to make strong progress on the research infrastructure goals and in the development and optimization of the methods needed for data analysis. We also have completed and published all of the data presented as preliminary data in the initial application. The clinical studies were held up by an unexpectedly long delay in obtaining final approval for our existing protocols - as noted by previous reviewers of our annual reports, this delay adversely affected the prospective study. Consistent with the recommendation of these prior reviewers, it was necessary to request a one-year no cost extension. This extension was formally requested and it has been approved, allowing us to continue the study and accrue additional clinical and microarray data. Overall, we believe that we have made good progress and continue to be productive in publishing the outcomes of this research and in advancing the scientific goals of this study.

REFERENCES

1 Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomized trials. Lancet, *351:* 1451-1467, 1998.

2 Early Breast Cancer Trialists Collaborative Group: Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. Lancet, *399:* 1-15, 1992.

3 Winer,E.P., Hudis,C., Burstein,H.J., Chlebowski,R.T., Ingle,J.N., Edge,S.B., Mamounas,E.P., Gralow,J., Goldstein,L.J., Pritchard,K.I., Braun,S., Cobleigh,M.A., Langer,A.S., Perotti,J., Powles,T.J., Whelan,T.J. and Browman,G.P. American Society of Clinical Oncology technology assessment on the use of aromatase inhibitors as adjuvant therapy for women with hormone receptor-positive breast cancer: status report 2002. J Clin Oncol, *20:* 3317-3327, 2002.

4 Ravdin,P. Aromatase inhibitors for the endocrine adjuvant treatment of breast cancer. Lancet, *359:* 2126-2127, 2002.

5 Bonneterre,J., Buzdar,A., Nabholtz,J.M., Robertson,J.F., Thurlimann,B., von Euler,M., Sahmoud,T., Webster,A. and Steinberg,M. Anastrozole is superior to tamoxifen as first-line therapy in hormone receptor positive advanced breast carcinoma. Cancer, *92:* 2247-2258, 2001.

6 Baum,M., Buzdar,A.U., Cuzick,J., Forbes,J., Houghton,J.H., Klijn,J.G., Sahmoud,T. and ATAC Trialists Group Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. Lancet, *359:* 2131-2139, 2002.

7 Ellis,M.J., Coop,A., Singh,B., Mauriac,L., Llombert-Cussac,A., Janicke,F., Miller,W.R., Evans,D.B., Dugan,M., Brady,C., Quebe-Fehling,E. and Borgs,M. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor- positive primary breast cancer: evidence from a phase III randomized trial. J Clin Oncol, *19:* 3808-3816, 2001.

8 Miller,W.R., Anderson,T.J. and Dixon,J.M. Anti-tumor effects of letrozole. Cancer Invest, *20 Suppl 2:* 15-21, 2002.

9 Smith,I.E. and Dowsett,M. Aromatase inhibitors in breast cancer. N Engl J Med, *348:* 2431-2442, 2003.

10 Clarke,R., Leonessa,F., Welch,J.N. and Skaar,T.C. Cellular and molecular pharmacology of antiestrogen action and resistance. Pharmacol Rev, *53:* 25-71, 2001.

11 Clarke,R. and Brünner,N. Acquired estrogen independence and antiestrogen resistance in breast cancer: estrogen receptor-driven phenotypes? Trends Endocrinol Metab, *7:* 25-35, 1996.

12 Clarke,R., Skaar,T.C., Bouker,K.B., Davis,N., Lee,Y.R., Welch,J.N. and Leonessa,F. Molecular and pharmacological aspects of antiestrogen resistance. J Steroid Biochem Mol Biol, *76:* 71-84, 2001.