JOURNAL OF CLINICAL ONCOLOGY

Randomized Phase II Neoadjuvant Comparison Between Letrozole, Anastrozole, and Exemestane for Postmenopausal Women With Estrogen Receptor–Rich Stage 2 to 3 Breast Cancer: Clinical and Biomarker Outcomes and Predictive Value of the Baseline PAM50-Based Intrinsic Subtype—ACOSOG Z1031

Matthew J. Ellis, Vera J. Suman, Jeremy Hoog, Li Lin, Jacqueline Snider, Aleix Prat, Joel S. Parker, Jingqin Luo, Katherine DeSchryver, D. Craig Allred, Laura J. Esserman, Gary W. Unzeitig, Julie Margenthaler, Gildy V. Babiera, P. Kelly Marcom, Joseph M. Guenther, Mark A. Watson, Marilyn Leitch, Kelly Hunt, and John A. Olson

From the Siteman Cancer Center; Washington University, St Louis, MO; American College of Surgeons Oncology Group Statistical Center, Mayo Clinic, Rochester, MN; Linberger Cancer Center, University of North Carolina, Chapel Hill; Duke University Cancer Center, Durham, NC; Helen Diller Cancer Center, University of California San Francisco, San Francisco, CA; Doctors Hospital of Laredo, Laredo; University of Texas MD Anderson Cancer Center, Houston; Simmons Cancer Center, University of Texas Southwestern, Dallas, TX; and Good Samaritan Hospital, Cincinnati, OH.

Corresponding author: Matthew J. Ellis, MB, BChir, PhD, Siteman Cancer Center, Washington University in St Louis, 660 South Euclid Ave, Campus Box 8056, St Louis, MO 63110; e-mail: mellis@dom.wustl.edu.

ABSTRACT

Purpose

Preoperative aromatase inhibitor (AI) treatment promotes breast-conserving surgery (BCS) for estrogen receptor (ER) –positive breast cancer. To study this treatment option, responses to three AIs were compared in a randomized phase II neoadjuvant trial designed to select agents for phase III investigations.

Patients and Methods

Three hundred seventy-seven postmenopausal women with clinical stage II to III ER-positive (Allred score 6-8) breast cancer were randomly assigned to receive neoadjuvant exemestane, letrozole, or anastrozole. The primary end point was clinical response. Secondary end points included BCS, Ki67 proliferation marker changes, the Preoperative Endocrine Prognostic Index (PEPI), and PAM50-based intrinsic subtype analysis.

Results

On the basis of clinical response rates, letrozole and anastrozole were selected for further investigation; however, no other differences in surgical outcome, PEPI score, or Ki67 suppression were detected. The BCS rate for mastectomy-only patients at presentation was 51%. PAM50 analysis identified AI-unresponsive nonluminal subtypes (human epidermal growth factor receptor 2 enriched or basal-like) in 3.3% of patients. Clinical response and surgical outcomes were similar in luminal A (LumA) versus luminal B tumors; however, a PEPI of 0 (best prognostic group) was highest in the LumA subset (27.1% v 10.7%; $P = .004$).

Conclusion

Neoadjuvant AI treatment markedly improved surgical outcomes. Ki67 and PEPI data demonstrated that the three agents tested are biologically equivalent and therefore likely to have similar adjuvant activities. LumA tumors were more likely to have favorable biomarker characteristics after treatment; however, occasional paradoxical increases in Ki67 (12% of tumors with > 5% increase after therapy) suggest treatment-resistant cells, present in some LumA tumors, can be detected by post-treatment profiling.

J Clin Oncol 29:2342-2349. © 2011 by American Society of Clinical Oncology

INTRODUCTION

Neoadjuvant aromatase inhibitor (AI) therapy is a low-toxicity approach that improves the breast conservation rate for postmenopausal women with clinical stage 2 or 3 estrogen receptor (ER) –positive breast cancer.[1] However, chemotherapy is still widely used in this setting despite promising small studies showing little advantage over an endocrine approach.[2] The lack of a practice standard reflects the absence of a phase III trial definitively comparing neoadjuvant aromatase inhibition with neoadjuvant chemotherapy. Unfortunately, the design of such a study is not straightforward, because pathologic complete response (CR) rates are low in ER-positive disease regardless of treatment modality, suggesting

other primary end points must be considered.[3,4] However, the clinical response rate (cRR) traditionally used in neoadjuvant endocrine studies is not verifiable, and radiologic response end points are not standardized.[5] Surgical outcome improvement is a logical primary end point, but surgical decisions are subject to bias when blinded treatments are not possible. The choice between anastrozole, letrozole, and exemestane is also an important consideration. The results of large randomized trials comparing these agents as adjuvant therapy are underway (eg, NCT00438529, NCT00248170, and NCT00541086). AIs have differences with respect to estradiol and estrone sulfate suppression, which could translate into differences in clinical responsiveness.[6]

An alternative to conventional primary end points for neoadjuvant endocrine therapy trials is the Ki67 proliferation biomarker.[7-10] Tumor Ki67 levels determined after initiation of neoadjuvant endocrine treatment are markedly more prognostic than baseline analysis,[11] and Ki67 data have been integrated into a post-treatment model that also includes pathologic stage and ER levels, referred to as the Preoperative Endocrine Prognostic Index (PEPI). Patients with pathologically node-negative T1 or T2 disease with a fully suppressed Ki67 level ($\leq 2.7\%$ or $\leq 1\%$ on a natural log scale) and persistent ER expression after completion of neoadjuvant endocrine therapy (PEPI of 0) were found to have such a low risk of relapse that adjuvant chemotherapy after neoadjuvant endocrine therapy may not be necessary.[12] Prognostic biomarker analysis based on tumor samples taken after treatment initiation is distinct from recent genomic approaches based on pretreatment sample analysis.[13-15] An integration of pre- and post-treatment biomarkers, could therefore improve prognostic algorithms and help identify patients for whom neoadjuvant endocrine therapy is appropriate, because adjuvant chemotherapy is unnecessary treatment.

To investigate these issues, the ACOSOG (American College of Surgeons Oncology Group) Z1031 trial, a randomized phase II study designed to select AIs for future studies, was conducted. The trial design prospectively incorporated an analysis of the PEPI approach (based on post-treatment sample analysis)[12] and the PAM50-based intrinsic subtype model (based on pretreatment sample analysis)[15] to compare these distinct approaches with prognostication for patients with ER-positive disease.

PATIENTS AND METHODS

Study Design

The Z1031 study is a phase II screening trial designed to determine which AIs (exemestane, letrozole, or anastrozole, one, two, or any of the three) should be chosen for future investigation. The study was supported by the Clinical Trials Support Unit and approved by the institutional review boards of all participating institutions; all patients provided signed informed consent. Eligible patients were menopausal (verified by bilateral surgical oophorectomy, no spontaneous menses for more than 1 year, or no menses for less than 1 year but follicle-stimulating hormone and estradiol levels in postmenopausal range), with clinical stage T2-T4c, N0-3, M0 breast cancer. Additional criteria were palpable lesion measuring more than 2 cm in at least one dimension; ER-positive disease with an Allred score of 6 to 8,[16] Eastern Cooperative Oncology Group performance status of 0 to 2, and a mammogram or ultrasound taken within 42 days before registration. Exclusion criteria included prior treatment for invasive breast cancer, hormone replacement therapy, megestrol acetate or raloxifene within 1 week before registration, inflamma-

tory breast cancer, and other malignancy less than 5 years before registration (except for nonmelanomatous skin cancer, lobular carcinoma in situ of the breast, contralateral ductal carcinoma in situ not treated with endocrine therapy, and carcinoma in situ of the cervix). At baseline, patients were categorized as marginal candidate for breast conservation, only eligible for mastectomy, or inoperable by standard mastectomy.

Treatment Administration and Evaluation

Eligible patients were treated with exemestane 25 mg daily, letrozole 2.5 mg daily, or anastrozole 1 mg daily for 16 to 18 weeks before surgery. Treatment was discontinued if disease progression was confirmed by ultrasound or mammogram, if severe toxicity (National Cancer Institute Common Terminology Criteria for Adverse Events version 3.0, grades 3 to 5) was reported as probably or definitely related to treatment, or if the patient withdrew consent. Within 14 days of registration, patients underwent a complete physical examination; assessment of Eastern Cooperative Oncology Group performance status; and tumor assessment by tape, ruler, or caliper. Patients were randomly assigned to treatment using a dynamic allocation procedure that balanced the marginal distributions of clinical tumor stage (T2 *v* T3 *v* T4) between treatment arms. Every 4 weeks, patients underwent a physical examination, toxicity assessment, and tumor assessment using WHO criteria. If tumor progression was suspected, ultrasound or mammogram was required for confirmation. Blood and biopsy specimens for correlative studies were collected at baseline and surgery or on discontinuation of endocrine therapy.

Statistical Considerations

The primary end point was the cRR (CRs and partial responses [PRs] as defined by WHO criteria) after 16 to 18 weeks of treatment. The last bidimensional clinical measurements, taken between 80 and 140 days after registration, were used to determine clinical tumor response. The trial was designed to select a treatment or treatments for further study so that there would be a 90% probability that the subset selected would include the AI with the truly highest cRR. Simulation studies with a sample size of 125 patients per arm generated a decision rule whereby the AI with the highest number of clinical responses (best agent) and any other agent(s) that had at most 12 fewer responses than the best agent would be included in the selected subset for further study. For example, the likelihood of selecting a single agent with the truly best response rate was 0.96 if the cRRs were 50%, 60%, and 80%, respectively. The likelihood of selecting two agents, if the cRRs for three agents were 50%, 70%, and 70%, respectively, was 0.90. An interval estimate for the true cRR for a given treatment was constructed using a 95% binomial CI for a single proportion. Pairwise differences between treatments in terms of true cRR were examined using 95% z-CIs for the true difference between independent proportions. Univariate logistic regression modeling was used to assess whether likelihood of a clinical response differed with respect to the baseline patient characteristics: histologic grade, cT stage, cN0, local laboratory, Allred score, or surgical status. Because none were found (likelihood ratio test *P* values > .05), assessing whether likelihood of clinical response differed with respect to treatment adjusting for significant baseline factors was abandoned. Among the subgroup of patients with PAM50 data, stratified logistic regression modeling was used to assess whether the likelihood of clinical response, breast-conserving surgery (BCS), or PEPI of 0 differed with respect to disease characteristics. The stratum was assigned treatment. For continuous and ordinal factors, Wilcoxon rank sum tests or Kruskal-Wallis tests were used to examine differences between intrinsic subtypes or treatment arms. Wilcoxon signed rank tests were used to examine pre- and post-treatment changes.

Biomarker Analysis

Tissue samples at baseline and surgery were submitted for PAM50-based intrinsic subtype assignment, ER Allred score, and Ki67 proliferation marker. Baseline tumor RNA was extracted from biopsies with at least 50% tumor; data were generated on an Agilent microarray platform (Santa Clara, CA) and analyzed using the PAM50 algorithm.[15] Array data normalization approaches and biomarker methodologies are provided in the Data Supplement. PAM50-based hierarchical clustering of Z1031 samples and subtype controls are provided in Appendix Figure A1, online only.
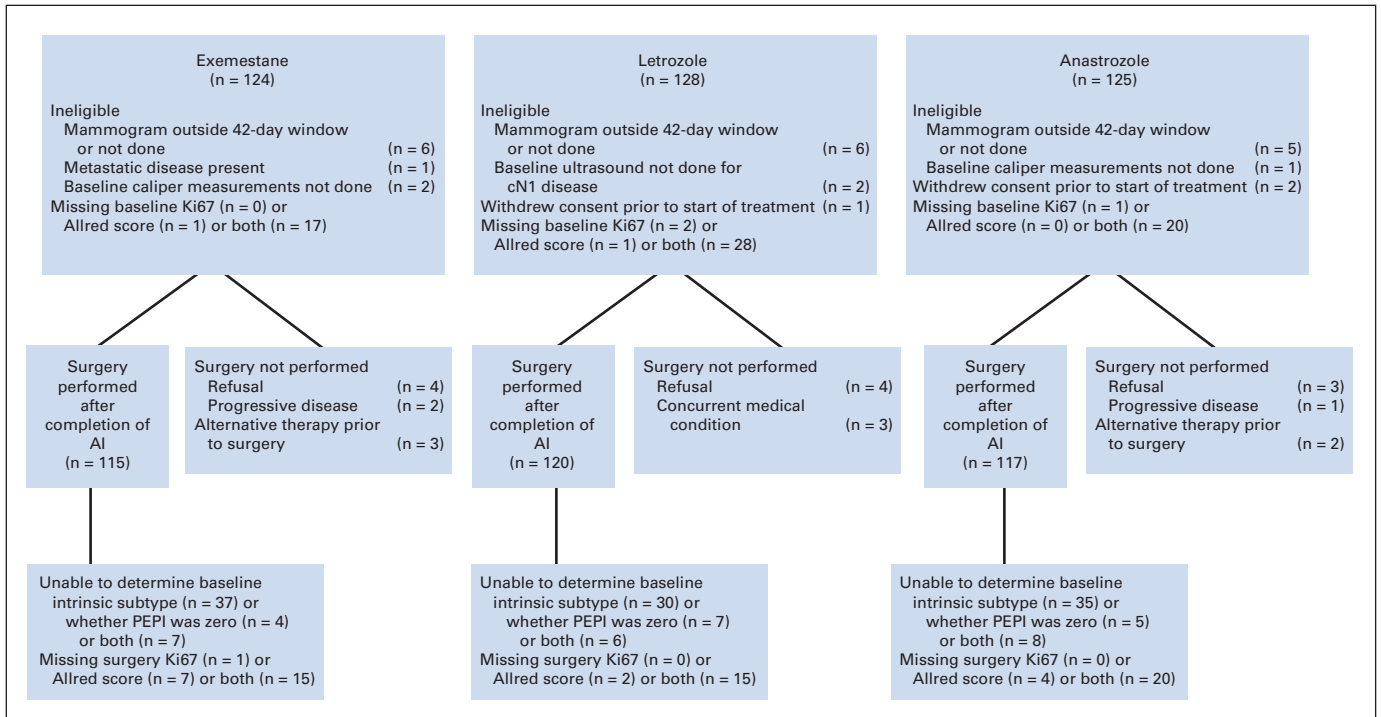
Fig 1. CONSORT diagram. AI, aromatase inhibitor; PEPI, Preoperative Endocrine Prognostic Index.

RESULTS

Patient Cohort

From January 4, 2006, to January 10, 2009, 377 postmenopausal women with clinical stage II or III ER-positive breast cancer were enrolled. One patient randomly assigned to receive letrozole and two patients randomly assigned to receive anastrozole withdrew consent before receiving any study treatment and were excluded. The remaining 374 patients compose the analysis cohort (Fig 1). Patient and disease characteristics are presented by treatment arm in Appendix Table A1, online only. A smaller percentage of patients randomly assigned to receive letrozole had cN0 disease, and a larger percentage of patients with an Allred score of 6 (by local testing) were randomly assigned to receive anastrozole. There was no difference in treatment duration between treatment arms (Appendix Table A2, online only).

Treatment Tolerability

No severe toxicity was reported by more than 5% of the patients. The most common grade 2 toxicity was hot flashes/flushes (Appendix Table A3, online only).

Clinical Response

Of the 124 patients receiving exemestane, 27 experienced CRs, 51 experienced PRs, 28 achieved stable disease (SD), and eight experienced progressive disease (PD)/sec primaries; baseline and/or final bidimensional measurements were not available (NA) for 10 patients because of treatment discontinuation (n = 5) or measurement failure (n = 5). Therefore, the cRR was 62.9% (95% CI, 53.8% to 71.4%). Of 127 patients receiving letrozole, 27 experienced CRs, 68 experienced PRs, 20 achieved SD, and six experienced PD; data for six were NA

because of treatment discontinuation (n = 3) or measurement failure (n = 3). The cRR was therefore 74.8% (95% CI, 66.3% to 82.1%). Of 123 patients receiving anastrozole, 22 experienced CRs, 63 experienced PRs, 20 achieved SD, and nine experienced PD; data for nine were NA because of treatment discontinuation (n = 2) and failure to measure (n = 7). Therefore, the cRR was 69.1% (95% CI, 60.1% to 77.1%). Letrozole exhibited the highest cRR, and the difference in the number of responses between letrozole and anastrozole was fewer than 12; thus, these two agents comprise the subset for further consideration (Table 1).

Table 1. Clinical Response Using WHO Criteria Based on ITT Population

Response	Exemestane (n = 124)		Letrozole (n = 127)		Anastrozole (n = 123)	
	No.	%	No.	%	No.	%
Clinical response at week 16 (WHO criteria with caliper measurements)						
Complete response	27	21.8	27	21.3	22	17.9
Partial response	51	41.1	68	53.5	63	51.2
No change	28	22.6	20	15.7	20	16.3
Disease progression	8	6.5	6	4.7	9	7.3
Off treatment because of toxicity/refusal	5	4.0	3	2.4	2	1.6
Measurements not done	5	4.0	3	2.4	7	5.7
ITT clinical response rate, %	62.9		74.8		69.1	
Range	53.8-71.4		66.3-82.1		60.1-77.1	
95% CI, % (difference in clinical response rates)						
Letrozole and exemestane			0.5 to 23.3			
Anastrozole and exemestane			−5.6 to 18.0			
Letrozole and anastrozole			−5.4 to 16.8			

Abbreviation: ITT, intent to treat.

Table 2. Surgical Procedures Performed After Neoadjuvant AI Therapy by Treatment Arm and Baseline Surgical Feasibility

Type of Surgical Candidate Before AI Therapy	Most Extensive Surgery Performed	Treatment Arm						Total	
		Exemestane (n = 115)		Letrozole (n = 120)		Anastrozole (n = 117)			
		No.	%	No.	%	No.	%	No.	%
Marginal for breast conservation (n = 189)	Breast conserved	52	45.2	48	40.0	57	48.7	157	83.1
	Mastectomy	9	7.8	14	11.8	9	7.7	32	16.9
Candidate for mastectomy only (n = 159)	Breast conserved	25	21.7	24	20.0	32	27.4	81	50.9
	Mastectomy	27	23.5	33	27.5	18	15.4	78	49.1
Inoperable by standard mastectomy (n = 4)	Breast conserved	1	0.9	1	0.8	1	0.9	3	75.0
	Mastectomy	1	0.9	—		—		1	25.0

Abbreviation: AI, aromatase inhibitor.

Surgical Outcomes

Seventeen patients did not undergo surgery because of patient refusal (n = 12), PD (n = 3), or concurrent medical conditions (n = 2). Another five patients received alternative systemic therapy before surgery. The remaining 352 patients form the surgical cohort (Fig 1). Fifty-one percent who were designated candidates for mastectomy only before therapy received BCS, and 83% of those considered marginal for BCS at baseline experienced successful breast conservation (Table 2). Pathologic findings were similar between treatment arms (Appendix Table A4, online only).

Analysis of Ki67 and PEPI by Treatment Arm

No differences were found between treatments with respect to baseline Ki67 levels (Kruskal-Wallis test P = .85) or changes in Ki67
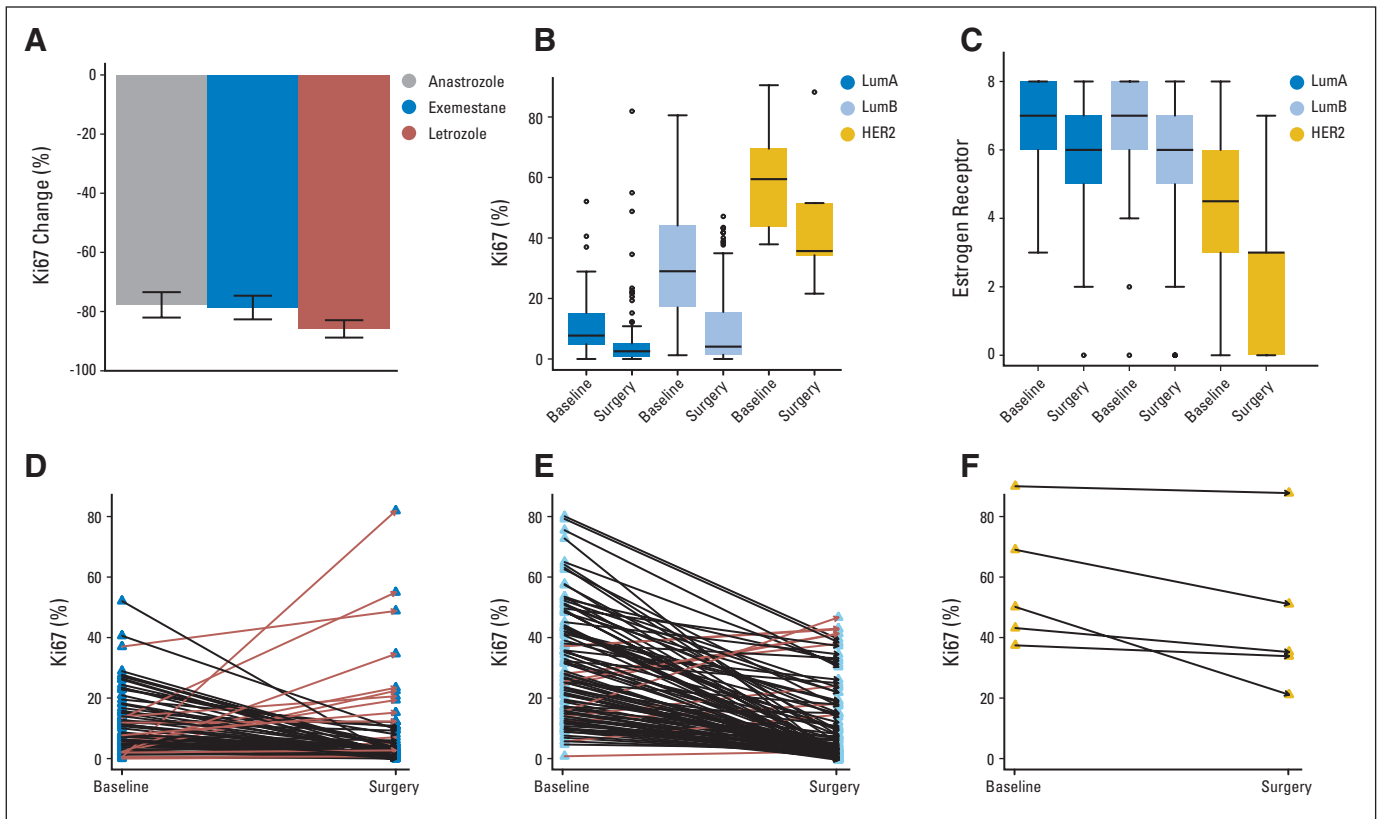

Fig 2. Distribution of biomarkers by treatment assignments and intrinsic subtypes. (A) Geometric mean percentage suppression of Ki67 from baseline by treatment arm to graphically summarize the overall treatment effect on each arm (Data Supplement). There was no significant difference in Ki67 suppression between treatment arms (Kruskal-Wallis P = .45, adjusted for three-way comparison). (B) Box plots of baseline and surgery Ki67 values in luminal A (LumA), luminal B (LumB), and human epidermal growth factor receptor 2 (HER2)–enriched breast cancer subtypes (P values listed in Table 3). (C) Box plots of baseline and surgery estrogen receptor Allred values in LumA, LumB, and HER2-enriched breast cancer subtypes (P values listed in Table 3). (D, E, F) Arrow plots showing individual paired Ki67 values in (D) LumA, (E) LumB, and (F) HER2-enriched breast cancer subtypes. Red arrows highlight patients with paradoxical increase in Ki67 values.

after treatment (Kruskal-Wallis test $P = .45$). Following the approach of Dowsett et al,[7] the geometric mean percentage change in Ki67 for each treatment was determined: anastrozole -78% (standard error of the mean [SEM], 4%), exemestane -81.2% (SEM, 3.5%), and letrozole -87.1% (SEM, 2.8%; Fig 2A). Within the limits of the sample size, the three treatments therefore had biologically equivalent effects. There were 315 patients in whom it could be determined whether the PEPI score was 0 or not (reasons for lack of data in the remaining patients included inadequate sample provision and absent axillary staging). No differences were found between the treatments with respect to the proportion of patients with PEPI of 0 ($P = .9$; anastrozole, 17.3%; letrozole, 15.9%; exemestane, 15.6%; Appendix Table A5, online only).

PAM50 Intrinsic Subtype, Ki67 Proliferation Index, and ER Allred Scores

Of the 315 patients with PEPI data, there were 102 patients for whom the PAM50 intrinsic subtype could not obtained, principally

because of inadequate frozen tumor tissue. The assignments were luminal A (LumA, 85 patients), luminal B (LumB, 119 patients), normal (two patients), human epidermal growth factor receptor 2 (HER2) enriched (six patients), and basal-like (one patient). The normal-like samples were not considered further, because these specimens may not have been sufficiently tumor rich to assign a breast cancer subtype.[15] Both the baseline Ki67 values and surgical specimen Ki67 values were significantly higher in LumB versus LumA breast cancer (Fig 2B; Table 3). The single basal-like patient demonstrated high pre- and post-treatment Ki67 values consistent with endocrine therapy resistance (38% and 26.8%, respectively) as well as low ER Allred scores (3 and 0, respectively). All five HER2-enriched patients had persistently high surgical Ki67 levels ($> 20\%$), consistent with high-level estrogen-independent growth (Table 3; Fig 2B, 2F). ER decreased after treatment in both LumA ($P = .001$) and LumB patients ($P \leq .001$; Table 3; Fig 2C). Paradoxical increases in Ki67 in the surgical sample compared with baseline greater than 5% were seen in

Table 3. ER and Ki67 Immunohistochemistry Biomarker Results in LumA Versus LumB Versus HER2-Enriched Breast Cancers Assigned by Microarray Profiling and Application of PAM50 Algorithm

Result	LumA	LumB	HER2 Enriched	P^*
Baseline Ki67				LumA v LumB: < .001
No.	89	119	6	
Median, %	7.7	29.0	59.4	
Range, %	0-52.1	1.2-80.5	37.8-90.5	
≤ 10%, %	57.3	6.7	0	LumA v HER2 enriched: .001
> 10%, %	42.7	93.3	100	LumB v HER2 enriched: .0032
Post-AI Ki67				
No.	82	114	5	
Median, %	2.6	4.1	35.7	LumA v LumB: .001
Range, %	0-81.9	0-47.1	21.6-88.2	
≤ 1%, %	32.9	16.7	0	LumA v HER2 enriched: < .001
> 1%, %	67.1	83.3	100	LumB v HER2 enriched: .0012
Change in Ki67 (after − before)				
No.	73	104	5	
Median, %	−4.6	−20.4	−8.0	LumA v LumB: < .001
Range, %	−51.5-81.4	−72.8-30.9	−29.0-−2.27	
Wilcoxon signed rank P	< .001	< .001	NS	LumA v HER2 enriched: NS
> 5% paradoxical increase, %	12.3	5.8	0	LumB v HER2 enriched: NS
Baseline Allred scores				
No.	90	118	6	
Median	7.0	7.0	4.5	LumA v LumB: NS
Range	3-8	0-8	0-8	
≤ 5, %	17.8	12.7	66.7	LumA v HER2 enriched: NS
≥ 6, %	82.2	87.3	33.3	LumB v HER2 enriched: NS
Post Allred scores				
No.	80	112	5	
Median	6.0	6	3	LumA v LumB: NS
Range	0-8	0-8	0-7	
≤ 5, %	35.0	38.4	80.0	LumA v HER2 enriched: NS
≥ 6, %	65.0	61.6	20.0	LumB v HER2 enriched: NS
Change in Allred Scores (before − after)				
No.	73	101	5	LumA v LumB: NS
Median	−1	−1	0	LumA v HER2 enriched: NS
Range	−7-4	−8-7	−4-2	
Wilcoxon signed rank P	.0012	< .001	NS	LumA v HER2 enriched: NS

NOTE. Nonparametric testing was used throughout to assign statistical significance because of non-normal distribution of Ki67 data.
Abbreviations: ER, estrogen receptor; LumA, luminal A; LumB, luminal B; HER2, human epidermal growth factor receptor 2; AI, aromatase inhibitor; NS, not significant.
*To account for multiple testing, $P < .01$ was considered significant for within-group comparisons of Ki67 and Allred score, and $P < .0055$ was considered significant for between-group comparisons of Ki67 and Allred score.

12.3% of LumA patients, 5.8% of LumB patients, and none of the HER2-enriched patients. Paradoxical Ki67 increases are highlighted with red arrow plots in Figures 2D to 2F for LumA, LumB, and HER2 enriched, respectively.

Clinical Response, Surgical Outcomes, and PEPI Scores in LumA and LumB Subtypes

No differences between the likelihood of BCS or clinical response were found between LumA and LumB tumors (Appendix Table A6, online only). However, 27.1% of patients with tumors considered LumA were assigned PEPI-0 status at surgery in contrast with 10.7% of patients assigned LumB status ($P = .004$; Appendix Table A7, online only). In univariable analysis, a baseline Ki-67 level less than or equal to 10% ($P = .018$) and LumA subtype status ($P = .004$) were significantly associated with an increased likelihood of a PEPI of 0. In multivariate analyses, LumA subtype assignment was the dominant factor predicting the likelihood of PEPI-0 status (Table 4).

DISCUSSION

The results from this study confirm earlier trial results demonstrating that marked improvements in surgical outcomes are achievable with neoadjuvant aromatase inhibitor therapy. This low toxicity approach is therefore a reasonable standard of care for selected patients with ER-rich HER2-negative breast cancer who desire breast conservation despite clinical stage 2 or 3 disease.[17,18] Limiting treatment to patients with tumor Allred scores of 6 or higher in Z1031 was based on of the findings of the letrozole P024 trial[19] and is a reasonable guideline for

clinical practice. The overall breast conservation rate of 68% in Z1031 compares well with the surgical outcome improvements achieved with doxorubicin and cyclophosphamide in the NSABP (National Surgical Adjuvant Breast and Bowel Project) B18 trial.[20]

The phase II statistical design selected letrozole or anastrozole for further study. However, no statistically significant differences with respect to Ki67 changes or PEPI values were detected. Thus, exemestane should not be necessarily excluded from investigation in the neoadjuvant setting, particularly if drug interactions or other factors preclude the use of nonsteroidal AIs. Importantly, the Ki67-based data demonstrate that the biologic activities of the three AIs studied are closely equivalent, therefore predicting similar activity as adjuvant therapies. In the case of exemestane and anastrozole, this has recently been shown to be the case. The NCIC (National Cancer Institute of Canada) MA27 trial showed no statistical differences in any trial end point in 7,576 patients who were randomly assigned to treatment with one or the other of these two agents.[21] It is therefore becoming increasingly illogical to embark on phase III adjuvant endocrine therapy studies on the basis of a superiority hypothesis without proof of greater biologic effectiveness of the experimental agent generated by a neoadjuvant Ki67 study. Although the neoadjuvant paradigm has been widely promoted for predicting the activity of adjuvant chemotherapy regimens,[22,23] our findings demonstrate that the case for endocrine agents also holds as long as conclusions are based on Ki67 suppression rather than pCR. The approximate 80% geometric mean Ki67 suppression observed with anastrozole in the IMPACT (Immediate Preoperative Anastrozole, Tamoxifen, or Combined with Tamoxifen) trial[7] is similar to our

			Wald Test	Univariate		Multivariate	
Factor	No.	PEPI Score of 0 (%)	P	OR	95% CI	OR	95% CI
Clinical T size							
T2	150	18.7	.433	1.41	0.60 to 3.30		
T3/T4a-c	56	14.3					
Clinical nodal status							
Positive	143	20.3	.144	1.94	0.80 to 4.70		
Negative	63	11.1					
HER2 (local)							
Negative	184	17.4	.575	1.54	0.34 to 7.05		
Positive	17	11.8					
Baseline Ki67, %							
≤ 10	50	28.0	.018	2.61	1.18 to 5.78		
> 10	135	12.6					
Baseline histologic grade							
1	35	20.0	.732	1.18	0.47 to 2.97		
2-3	170	17.1					
Baseline ER score, Allred (central testing)							
6-8	159	16.4	.924	1.06	0.34 to 3.32		
≤ 5	26	15.4					
Intrinsic subtype							
LumA	85	27.1	.004	2.99	1.42 to 6.31	2.99	1.42 to 6.31
LumB	121	10.7					

Table 4. Results of Stratified Logistic Regression Modeling of Baseline Factors to Potential to Predict PEPI-0 Status

NOTE. Aromatase inhibitor treatment was stratification factor. Six tumors were assigned nonluminal intrinsic subtypes: five HER2 enriched and one basal-like. None of these patients achieved PEPI-0 status and were therefore not included in analysis, which dichotomized tumors as either LumA or LumB according to PAM50 intrinsic subtype. This analysis included patients who could be assigned to non–PEPI-0 group, even though some elements of analysis were not available.
Abbreviations: PEPI, Preoperative Endocrine Prognostic Index; OR, odds ratio; HER2, human epidermal growth factor receptor 2; ER, estrogen receptor; LumA, luminal A; LumB, luminal B.

observations in Z1031 (Fig 2A), underscoring the consistency of the Ki67 approach.

In terms of the identification of suitable patient populations for neoadjuvant endocrine therapy, PAM50 intrinsic subtype analysis seems useful for excluding uncommon nonluminal intrinsically endocrine therapy–resistant tumors present, despite trial eligibility requiring strong ER expression (3.3% were HER2 enriched or basal-like). Both LumA and LumB tumors were highly endocrine-therapy responsive at the clinical and biomarker level; however, LumB tumors had significantly higher post-treatment Ki67 levels, consistent with the worse prognosis associated with this subtype.[14,15] Therefore, patients with LumA breast cancer may be particularly suitable for neoadjuvant endocrine therapy, because approximately one quarter of these patients have a PEPI of 0, implying that a prognosis with adjuvant endocrine therapy may be sufficient without adjuvant chemotherapy treatment.[12] In the absence of a PAM50 result, a baseline Ki67 level of 10% or less was significant in univariable analysis for predicting PEPI-0 status and could also be used to identify patients at baseline suitable for neoadjuvant endocrine therapy. However, Ki67 produced a smaller LumA-like population (27%) than the gene expression profiling–based LumA definition (41%). Practically speaking, the Ki67 assay provides a backup approach when tissue sampling is inadequate for gene expression profiling, and the two tests can be considered complementary.

Paradoxical Ki67 increases after neoadjuvant endocrine therapy have been observed in other studies.[7,8] In Z1031, paradoxical rises were observed even when both the baseline Ki67 was low and the subtype assignment was LumA (Fig 2D), suggesting the increase was not simply the result of erroneously low baseline Ki67 values. A possible explanation is that some tumors assigned good-prognosis status harbor occult treatment-resistant cellular populations that become evident after neoadjuvant endocrine treatment. Decreased levels of ER after treatment, as seen in this study as well as in previously published work,[8,12] also likely reflect the biology of endocrine resistance. In general, tumor heterogeneity is a likely explanation for why baseline tests, whether simple Ki67 or complex gene expression analysis, fail to identify ultra–low-risk populations in the setting of stages 2 and 3 disease.[14,24] Neoadjuvant endocrine therapy and repeated tumor sampling is a practical approach to address the heterogeneity issue.

Conclusions regarding the routine use of neoadjuvant endocrine therapy from this trial will be strengthened by relapse-free survival data. The study design requires that each patient is observed for 10 years. Of particular interest will be the prognostic integrity of the PEPI score, the PAM50 subtype assignments, and the in-breast recurrence rate for patients who experience neoadjuvant AI-assisted breast conservation. The development of a reliable and controlled clinical testing environment for both the PAM50 test as well as the Ki67 assay is also critical for further progress, particularly given the difficulties of methodologically reproducing data with gene expression signatures in the research setting, even when the genomic algorithm is fixed.[25,26]

AUTHOR CONTRIBUTIONS

Conception and design: Matthew J. Ellis, Marilyn Leitch, Kelly Hunt, John A. Olson
Administrative support: Matthew J. Ellis, Kelly Hunt
Provision of study materials or patients: Matthew J. Ellis, Laura J. Esserman, Gary W. Unzeitig, Julie Margenthaler, Gildy V. Babiera, P. Kelly Marcom, Joseph M. Guenther, Marilyn Leitch, Kelly Hunt
Collection and assembly of data: Matthew J. Ellis, Vera J. Suman, Jeremy Hoog, Li Lin, Jacqueline Snider, Joel S. Parker, Katherine DeSchryver, D. Craig Allred, Laura J. Esserman, Gary W. Unzeitig, Gildy V. Babiera, P. Kelly Marcom, Joseph M. Guenther, Mark A. Watson, Marilyn Leitch, Kelly Hunt, John A. Olson
Data analysis and interpretation: Matthew J. Ellis, Vera J. Suman, Aleix Prat, Joel S. Parker, Jingqin Luo, D. Craig Allred, Kelly Hunt, John A. Olson
Manuscript writing: Matthew J. Ellis, Vera J. Suman, Jeremy Hoog, Li Lin, Jacqueline Snider, Aleix Prat, Joel S. Parker, Jingqin Luo, Katherine DeSchryver, D. Craig Allred, Laura J. Esserman, Gary W. Unzeitig, Julie Margenthaler, Gildy V. Babiera, P. Kelly Marcom, Joseph M. Guenther, Mark A. Watson, Marilyn Leitch, Kelly Hunt, John A. Olson
Final approval of manuscript: Matthew J. Ellis, Vera J. Suman, Jeremy Hoog, Li Lin, Jacqueline Snider, Aleix Prat, Joel S. Parker, Jingqin Luo, Katherine DeSchryver, D. Craig Allred, Laura J. Esserman, Gary W. Unzeitig, Julie Margenthaler, Gildy V. Babiera, P. Kelly Marcom, Joseph M. Guenther, Mark A. Watson, Marilyn Leitch, Kelly Hunt, John A. Olson

REFERENCES

1. Chia YH, Ellis MJ, Ma CX: Neoadjuvant endocrine therapy in primary breast cancer: Indications and use as a research tool. Br J Cancer 103:759-764, 2010

2. Semiglazov VF, Semiglazov VV, Dashyan GA, et al: Phase 2 randomized trial of primary endocrine therapy versus chemotherapy in postmenopausal patients with estrogen receptor-positive breast cancer. Cancer 110:244-254, 2007

3. Ellis MJ: Neoadjuvant endocrine therapy for breast cancer: More questions than answers. J Clin Oncol 23:4842-4844, 2005

4. Tan MC, Al Mushawah F, Gao F, et al: Predictors of complete pathological response after neoadjuvant systemic therapy for breast cancer. Am J Surg 198:520-525, 2009

5. Olson JA Jr, Budd GT, Carey LA, et al: Improved surgical outcomes for breast cancer patients receiving neoadjuvant aromatase inhibitor therapy: Results from a multicenter phase II trial. J Am Coll Surg 208:906-914, 2009; discussion, 915-916

6. Dixon JM, Renshaw L, Young O, et al: Letrozole suppresses plasma estradiol and estrone sulphate more completely than anastrozole in postmenopausal women with breast cancer. J Clin Oncol 26:1671-1676, 2008

7. Dowsett M, Smith IE, Ebbs SR, et al: Short-term changes in Ki-67 during neoadjuvant treatment of primary breast cancer with anastrozole or tamoxifen alone or combined correlate with recurrence-free survival. Clin Cancer Res 11:951s-958s, 2005

8. Ellis MJ, Coop A, Singh B, et al: Letrozole inhibits tumor proliferation more effectively than tamoxifen independent of HER1/2 expression status. Cancer Res 63:6523-6531, 2003

9. Baum M, Budzar AU, Cuzick J, et al: 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

10. Thurlimann B, Keshaviah A, Coates AS, et al: A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. N Engl J Med 353:2747-2757, 2005

11. Dowsett M, Smith IE, Ebbs SR, et al: Prognostic value of Ki67 expression after short-term presurgical endocrine therapy for primary breast cancer. J Natl Cancer Inst 99:167-170, 2007

12. Ellis MJ, Tao Y, Luo J, et al: Outcome prediction for estrogen receptor-positive breast cancer based on postneoadjuvant endocrine therapy tumor characteristics. J Natl Cancer Inst 100:1380-1388, 2008

13. Paik S, Shak S, Tang G, et al: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351: 2817-2826, 2004

14. Nielsen TO, Parker JS, Leung S, et al: A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clin Cancer Res 16:5222-5232, 2010

15. Parker JS, Mullins M, Cheang MC, et al: Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 27:1160-1167, 2009

16. Allred DC, Harvey JM, Berardo M, et al: Prognostic and predictive factors in breast cancer by immunohistochemical analysis. Mod Pathol 11:155-168, 1998

17. Eiermann W, Paepke S, Appfelstaedt J, et al: Preoperative treatment of postmenopausal breast cancer patients with letrozole: A randomized double-blind multicenter study. Ann Oncol 12:1527-1532, 2001

18. Smith IE, Dowsett M, Ebbs SR, et al: Neoadjuvant treatment of postmenopausal breast cancer with anastrozole, tamoxifen, or both in combination: The Immediate Preoperative Anastrozole, Tamoxifen, or Combined with Tamoxifen (IMPACT) multicenter double-blind randomized trial. J Clin Oncol 23:5108-5116, 2005

19. Ellis MJ, Coop A, Singh B, et al: 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

20. Fisher B, Brown A, Mamounas E, et al: Effect of preoperative chemotherapy on local-regional disease in women with operable breast cancer: Findings from National Surgical Adjuvant Breast and Bowel Project B-18. J Clin Oncol 15:2483-2493, 1997

21. Goss PE, Ingle JN, Chapman J-AW, et al: Final analysis of NCIC CTG MA 27: A randomized phase III trial of exemestane versus anastrozole in postmenopausal women with hormone receptor positive breast cancer. Cancer Res 70:75s, 2010 (suppl 24; abstr S1-1)

22. Pusztai L: Preoperative systemic chemotherapy and pathologic assessment of response. Pathol Oncol Res 14:169-171, 2008

23. Symmans WF, Peintinger F, Hatzis C, et al: Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol 25:4414-4422, 2007

24. Albain KS, Barlow WE, Shak S, et al: Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: A retrospective analysis of a randomised trial. Lancet Oncol 11:55-65, 2010

25. Weigelt B, Mackay A, A'Hern R, et al: Breast cancer molecular profiling with single sample predictors: A retrospective analysis. Lancet Oncol 11: 339-349, 2010

26. Perou CM, Parker JS, Prat A, et al: Clinical implementation of the intrinsic subtypes of breast cancer. Lancet Oncol 11:718-719, 2010; author reply, 720-721

■ ■ ■