ACOSOG Z1031, a randomized phase 2 neoadjuvant comparison between letrozole, anastrozole and exemestane for postmenopausal women with ER rich stage 2/3 breast cancer: Clinical and Biomarker outcomes.

Ellis et al

Data Supplement 2

RNA Preparation Methodology: Frozen biopsies were sectioned, H/E stained, and reviewed by a pathologist to determine tumor cellularity. If >=50% of the cells within the section are tumor cells then 50um sections were sliced and placed into Trizol (Invitrogen) to be homogenized for RNA isolation. Total RNA was extracted from the homogenates with chloroform, precipitated with Isopropanol, washed with 70% Ethanol, and eluted in Milli-Q water. The RNA was then treated with DNase (Qiagen) and purified using the RNeasy MinElute Cleanup Kit (Qiagen). Purified total RNA samples were then assayed for quality using either an Agilent Bioanalyzer or a BioRad Experion.

Microarray data generation: All fresh frozen human tumor and normal tissue samples were collected using Institutional Review Board (IRB) approved protocols. Total RNA was purified, amplified and labeled (Cy5-sample, Cy3-control), and microarray hybridizations were performed using Agilent 4X44K Whole Human Genome Microarrays. For Cy3-controls, we used Stratagene Human Universal Reference[1] enriched with equal amounts of RNA from the

MCF7 and ME16C cell lines. Microarrays were hybridized overnight, washed, dried, and scanned using an Agilent Scanner. The image files were analyzed and loaded into the UNC-CH Microarray Database (https://genome.unc.edu/). Final normalized log2 ratios (Cy5-sample/Cy3-control) for each probe were obtained after removing probes with a lowess normalized intensity value <10 in the Cy5-sample and/or the Cy3-control.

Intrinsic molecular subtyping of the Z1301 microarray data set. The main molecular entities of breast cancer (Luminal A, Luminal B, HER2-enriched, Basal-like and Normal Breast-like) were identified using the latest intrinsic subtype classification, which is the PAM50 microarray-based assay described in Parker et al.[2] However, to account for potential technical biases such as differences in platforms and laboratory protocols between the Z1031 data set and the PAM50 training set[2], we normalized the Z1301 data set before applying the PAM50 algorithm. This is a critical step that must be taken into consideration for appropriate subtype predictions as our group and others have previously shown[3,4]. To appropriately normalize the Z1301 microarray data, which is entirely composed of ER-positive tumors, we needed to estimate technical bias from a cohort with equivalent subtype representation as the PAM50 training set. A set including ER-negative, HER2-positive and true normal breast samples were therefore profiled under the same platform, protocol and laboratory (i.e. Wash University) as the Z1301 data set. A total of 40 prototypic samples of the various molecular subtypes were selected to match the distribution in the PAM50 training set. The mean expression of each PAM50 gene was calculated in this set of 40.

These values obtained from the prototypical sample set were used for calibration purposes and were included as a new *WashU_4X44K_Calibration* column in the *CalibrationParameter* file, which has been already provided in the PAM50 algorithm[2]. This adjustment step assures that each sample is placed in the appropriate 'prototypic space' by subtracting the calibration value of each PAM50 gene from the expression value of the corresponding gene in each tested sample. Finally, single sample predictions were performed in every Z1301 sample using the *WashU_4X44K_Calibration* parameter. The 40 microarray data sets that provide the intrinsic prototypes for centroid-based subtype assignments can be retrieved via the GEO accession number GSE26082. The linked PAM50 gene expression data and the Ki67 and ER data used to generate Table 4 are provided in the supplemental data spread sheet.

Immunohistochemistry Methods for Estrogen Receptor (ER) and Ki67:

Briefly, the antibody for Ki67 (clone Sp6, Neomarkers) and ER (clone 6F-11, Novocatra) was applied at a 1:200 dilution for 1 hour after antigen retrieval in citrate buffer pH 6.0 at 95C in a steamer. Secondary antibody, species specific, was performed using an HRP labeled polymer which is conjugated with secondary antibodies (Envision–HRP, Dako). Staining was completed with 3'3'diaminobenzidine as the chromagen. Counter staining was accomplished using Mayer's Hematoxylin (Biogenex). Ki67 Scoring was conducted by taking three separate images of tumor rich areas using an Olympus DP 70 camera by

scanning over the whole slide at 40X magnification. These images were printed out onto a square 3/8 “X 3/8” grid and point counted according to the following schema: Reader 1 counted tumor cells and Ki67+ cells that intersect with first grid line. This process was repeated on every third gridline until one of the following criteria is met:

1) Cells were counted from three separate, 40X images, or
2) All the cells on the slide are counted if three images cannot be obtained.

Reader two counted tumor cells and Ki67+ cells that intersect with a grid line different than the one which reader one started with. The number of Ki67 positive cells and total number of tumor cell counts are combined from reader 1 and reader 2 and an average percent positive Ki67 value was obtained. ER Scoring was accomplished by using a surgical pathologist who used established criteria (Allred Score).

Calculation of the geometric mean and confidence interval

The values in Ki67 were replaced by 0.25 and 0.05 at baseline and Surgery respectively. The geometric mean was calculated as the exponentiated arithmetic mean of measurements in the logarithm scale whereas the standard error of mean (SEM) was approximated, due to asymmetric CIs, by dividing the corresponding 95% confidence interval length by 4.

References

1. Novoradovskaya N, Whitfield M, Basehore L, et al: Universal Reference RNA as a standard for microarray experiments. BMC Genomics 5:20, 2004

2. Parker JS, Mullins M, Cheang MCU, et al: Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. J Clin Oncol 27:1160-1167, 2009

3. Perou C, Parker J, Prat A, et al: Clinical implementation of the intrinsic subtypes of breast cancer. Lancet Oncol 11:718-719, 2010

4. Lusa L, McShane LM, Reid JF, et al: Challenges in Projecting Clustering Results Across Gene Expression Profiling Datasets. J. Natl. Cancer Inst. 99:1715-1723, 2007

Treatment Course and Toxicity Summary

Among the 124 women randomized to exemestane who began treatment, 110 (88.7%) completed treatment without a disease event. Two patients discontinued treatment due to withdrawal of consent and 4 patients discontinued due excessive toxicity (including grade 3 confusion/psychosis (1 pt), persistent grade 3 diarrhea (1 pt), grade 3 nausea and vomiting (1 pt), and grade 3 allergic reaction (1pt). Eight patients had disease progression - local /regional disease progression (5 pts); metastatic disease progression (2 pts); and second primary

disease (1 pt). The median duration of treatment among these 124 patients was 16 weeks (range: 2 – 23 weeks).

Among the 127 treated with letrozole, 116 (91.3%) completed treatment without a disease event. Three patients discontinued treatment due to withdrawal of consent, 2 patients discontinued due to excessive toxicity (including grade 3 confusion/hematoma/seizure (1 pt) and grade 3 muscle weakness (1 pt)) and six patients had disease progression - local /regional disease progression (4 pts); metastatic disease progression (1 pt); and second primary disease (1 pt). The median duration of treatment among these 127 patients was 16 weeks (range: 2 – 21 weeks).

Among the 123 treated with anastrozole 112 (91.1%) completed treatment without a disease event. Two patients discontinued treatment due to refusal and nine patients had disease progression - local /regional disease progression (7 pts) and second primary disease (2 pts). The median duration of treatment among these 123 patients was 16 weeks (range: < 1 – 25 weeks).