Clinical Cancer Research

Estrogen Receptor Binding (^{18}F-FES PET) and Glycolytic Activity (^{18}F-FDG PET) Predict Progression-Free Survival on Endocrine Therapy in Patients with ER$^+$ Breast Cancer

Brenda F. Kurland[1], Lanell M. Peterson[2], Jean H. Lee[2], Erin K. Schubert[3], Erin R. Currin[4], Jeanne M. Link[5], Kenneth A. Krohn[2], David A. Mankoff[3], and Hannah M. Linden[4]

Abstract

Purpose: ^{18}F-fluoroestradiol (FES) PET scans measure regional estrogen binding, and ^{18}F-fluorodeoxyglucose (FDG) PET measures tumor glycolytic activity. We examined quantitative and qualitative imaging biomarkers of progression-free survival (PFS) in breast cancer patients receiving endocrine therapy.

Experimental Design: Ninety patients with breast cancer from an estrogen receptor–positive (ER$^+$), HER2$^-$ primary tumor underwent FES PET and FDG PET scans prior to endocrine therapy (63% aromatase inhibitor, 22% aromatase inhibitor and fulvestrant, 15% other). Eighty-four had evaluable data for PFS prediction.

Results: Recursive partitioning with 5-fold internal cross-validation used both FES PET and FDG PET measures to classify patients into three distinct response groups. FDG PET identified 24 patients (29%) with low FDG uptake, suggesting indolent tumors. These patients had a median PFS of 26.1 months (95% confidence interval, 11.2–49.7). Of patients with more FDG-avid tumors, 50 (59%) had high average FES uptake, and 10 (12%) had low average FES uptake. These groups had median PFS of 7.9 (5.6–11.8) and 3.3 months (1.4–not evaluable), respectively. Patient and tumor features did not replace or improve the PET measures' prediction of PFS. Prespecified endocrine resistance classifiers identified in smaller cohorts did not individually predict PFS.

Conclusions: A wide range of therapy regimens are available for treatment of ER$^+$ metastatic breast cancer, but no guidelines are established for sequencing these therapies. FDG PET and FES PET may help guide the timing of endocrine therapy and selection of targeted and/or cytotoxic chemotherapy. A multicenter trial is ongoing for external validation. *Clin Cancer Res; 23(2); 407–15. ©2016 AACR.*

Introduction

Patients with distant metastatic breast cancer from an estrogen receptor–positive (ER$^+$) primary tumor are rarely cured but often live for many years with their disease (1). Clinical management of these patients is not standardized but generally involves a series of endocrine therapy treatment regimens, cytotoxic chemotherapy, and additional targeted agents (2, 3). Many patients achieve a durable response to endocrine therapy with few toxicities; however, not all patients benefit from endocrine therapy, and most responding tumors eventually become refractory (4). The primary biomarkers for directing therapy for these patients are *in vitro* assays of hormone receptor [ER, progesterone receptor (PgR)] expression. Beyond binary ER classification, higher levels of tumor ER expression are associated with greater clinical benefit from endocrine therapy (5).

In vivo imaging using ^{18}F-fluoroestradiol (FES) PET measures regional estrogen binding in breast cancer tumors (6), and FES uptake quantitation has good agreement with ER expression measured by IHC (7, 8). Early clinical studies summarized FES uptake in a single field of view (FOV) to predict clinical response to endocrine therapy (9–11). We extend that approach by using FES PET imaging to evaluate quantitative and qualitative ER expression at multiple disease sites in a torso survey performed immediately following the single FOV scan and by including FDG uptake to distinguish between indolent tumors and those that are metabolically aggressive (12). Prior clinical studies with FES PET imaging prior to endocrine therapy showed promising results, even with small sample sizes and limited follow-up (9–11). Our primary goal for this work was to assess imaging classifiers suggested by prior research as predictors of response to endocrine therapy in a

[1]Department of Biostatistics, University of Pittsburgh, Pittsburgh, Pennsylvania. [2]Department of Radiology, University of Washington, Seattle, Washington. [3]Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania. [4]Division of Medical Oncology, University of Washington, Seattle, Washington. [5]Department of Diagnostic Radiology, Oregon Health & Science University, Portland, Oregon.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Prior presentation: This study was presented in part at the Society of Nuclear Medicine and Molecular Imaging Annual Meeting, Vancouver, BC, Canada, June 8–12, 2013.

Corresponding Author: Brenda F. Kurland, University of Pittsburgh, Suite 325 Sterling Plaza, 201 North Craig St, Pittsburgh, PA 15213. Phone: 412-383-1128; Fax: 412-383-1535; E-mail: bfk10@pitt.edu

doi: 10.1158/1078-0432.CCR-16-0362

Translational Relevance

Currently, a wide range of therapy regimens are available for treatment of estrogen receptor–positive (ER^+) metastatic breast cancer, including endocrine therapy, cytotoxic chemotherapy, and molecularly targeted therapy. However, no guidelines are established for sequencing or combining these therapies. ^{18}F-fluorodeoxyglucose (FDG) PET and ^{18}F-fluoroestradiol (FES) PET may help guide the timing of endocrine therapy and selection of other targeted and/or cytotoxic chemotherapy. FDG PET is widely used to measure tumor glycolytic activity. FES PET measures regional estrogen binding and is not yet available clinically. Unlike tissue sampling, these imaging assays may evaluate the entire burden of metastatic disease. By identifying indolent tumors and tumors likely to respond to endocrine therapy, FDG PET and FES PET may be developed as predictive biomarkers to inform decisions about when to prescribe toxic and/or costly therapies as part of the often lengthy sequence of treatments for ER^+ metastatic breast cancer.

larger cohort with mature time-to-event endpoints. We explore these imaging biomarkers and patient/disease characteristics as predictors of progression-free survival (PFS), overall survival (OS), and clinical benefit (6+ months on endocrine therapy without progressive disease).

Materials and Methods

Patients

Study participants were recruited from the University of Washington Medical Center and the Seattle Cancer Care Alliance (UWMC/SCCA; Seattle, WA). Eligible patients were planning endocrine therapy for primary, recurrent, or metastatic breast cancer and had an ER-expressing primary tumor. Each patient provided informed consent for one of four prospective, observational FES PET studies evaluating: (i) heterogeneity of FES uptake in patients with advanced breast cancer; (ii) agreement of FES uptake to *in vitro* assay of ER, (iii) FES uptake as a predictor of therapy response; or (iv) patients with newly diagnosed stage IV disease to be treated with endocrine therapy. Use of FES was carried out under the University of Washington Radioactive Drug Research Committee (i–iii) or NCI IND #79005 approval (iv).

Enrolled patients were planning to receive one of the following standard-of-care treatments: tamoxifen ± ovarian suppression, aromatase inhibitor (AI) ± fulvestrant (with ovarian suppression in premenopausal patients), or fulvestrant or ovarian suppression alone. A washout period of 6 to 8 weeks for tamoxifen or fulvestrant was required prior to FES PET imaging, as these agents block the tracer from binding to the receptor. No washout period was required for AIs. Some patients were switching from a nonsteroidal to a steroidal AI due to a lack of response or disease progression following initial response or clinical benefit. In other cases, AI was continued and fulvestrant was added. Patient treatment was selected in advance of FES PET; referring physicians had access to imaging results, but FES PET was not used to make treatment decisions.

In vitro assays

Primary tumor IHC results were obtained by chart review when possible. For samples analyzed at the UWMC/SCCA, positive staining in >5% cells indicated a positive result. Slides or tissue blocks were requested for reanalysis when necessary.

PET imaging

Patients underwent FDG and FES studies prior to or shortly following initiation of or changes to endocrine therapy. Imaging was performed with a GE ADVANCE PET or GE Discovery STE PET/CT scanner operating in the high-sensitivity mode. The two scanners have nearly identical PET components, providing comparable quantitative PET data. For the 90 patients in the primary analysis, all but 3 FES scans were performed using the GE ADVANCE PET scanner; FDG scans for those 3 patients and an additional 12 were performed on the DSTE PET/CT. Rigorous cross-calibration was performed between the two PET scanners, for overall scale factor and resolution matching, using the NEMA-NU2 image quality phantom (13).

FDG imaging

2-[^{18}F-]fluoro-2-deoxy-D-glucose (FDG) was produced on site by University of Washington radiochemists according to established techniques, with quality control testing done on every batch (14). FDG imaging was done according to standard clinical procedures. Briefly, regardless of which machine was used, a dose of 260 to 370 MBq (7–10 mCi) of FDG was administered intravenously following a minimum 4-hour fast. Patients rested comfortably for 45 to 60 minutes in a supine position prior to FDG PET scanning. For the PET/CT device, a low-dose CT scan provided data for attenuation correction. On the ADVANCE, 3-minute attenuation scans were constructed for each of 5 FOVs. Each emission scan was 7 minutes per FOV. Attenuation or CT-corrected emission data was reconstructed by standard filtered back-projection after correction for scattered and random coincidences. Reconstruction used a 10-mm Hanning Filter producing a 128 pixel × 128 pixel × 35 slice imaging volume, yielding a reconstructed in-plane spatial resolution of 4.29 mm. A subset of scans was also reconstructed using iterative methods (ordered subset expectation maximization); comparisons of SUV quantification are described in Supplementary Data.

FES imaging

16α-^{18}F-fluoro-17β-estradiol (FES) was prepared locally by the PET Radiochemistry Group at the University of Washington (Seattle, WA) using established methods for synthesis and quality assurance. All tracer characteristics, including specific activity and the total mass of FES in the injectate, were well within established ranges, as described previously (15, 16). Per protocol, following a 25-minute attenuation scan (or low-dose CT scan on the PET/CT), a dose of 148 to 222 MBq (4–6 mCi) FES was infused intravenously in a volume of 20 mL isotonic PBS containing less than 15% of ethanol by volume over 2 minutes, and dynamic imaging was performed for 60 minutes over a 15-cm body region containing the most prominent tumor sites, starting at the beginning of infusion. Following the dynamic data collection, a neck-to-pelvis torso survey covering 5 × 15 cm axial FOVs was performed using 5-minute emission and 3-minute postinjection attenuation scans (or low-dose CT on the PET/CT) per FOV. Although dynamic scans were performed to

maintain a consistent protocol, ongoing analysis of FES data showed that static torso scans were sufficient for measuring FES uptake and allowed analysis of total disease burden. Therefore, static measurements from a torso sweep have become the focus of future analysis. Attenuation or CT-corrected emission data were reconstructed according to the same specifications as the FDG scans, as described above.

FDG PET and PET/CT image analysis

FDG images were processed as for a typical clinical scan, corrected for radioactive decay of the tracer, and normalized to the injected dose (ID) and body weight (BW). This results in regional standardized uptake values (SUV): SUV = A/(ID/BW), where A is the tissue tracer uptake in microcuries per gram for the hottest pixel in the tumor (SUV_{max}); ID is the injected dose in millicuries, and BW is the body weight in kilograms. SUL_{max} was also calculated, normalized by lean body mass (LBM) instead of BW. Lesions identified on FDG PET were corroborated by CT and/or other imaging.

FES PET qualitative assessment. Qualitative (visual) assessment of FES was performed using attenuation-corrected torso images. Active sites of disease visualized on the clinical FDG PET scan or CT scan were matched to the FES torso scan using a rigid coregistration/fusion module (PMOD Technologies Ltd.). Each disease site was evaluated as positive or negative for FES uptake, compared with normal tissue background. These qualitative assessments were summarized at the patient level as presence/absence of any site(s) qualitatively negative for FES uptake. Qualitative assessments were conducted by a single experienced observer with simultaneous access to both FES and FDG images. Three experienced readers evaluated a subset of 19 cases, with no disagreements on scan-level qualitative assessment (16).

FES PET quantitative assessment. For quantitative assessment of FES uptake, lesions were identified and localized using the FDG scan matched to the FES scan, as described above. A cubic region of interest was drawn on three adjacent planes (approximately 3 cc total volume) to identify the area with the greatest FES uptake for each disease site. Up to three lesions were quantified on the summed dynamic scan, and all lesions recorded on the clinical FDG scan (other than liver lesions) were quantified on the static torso survey. Partial-volume correction was not applied because only sites 1.5 cm^3 or greater were included. FES uptake was quantified as the standardized uptake value (SUV, or SUL with LBM correction) on both the 30- to 60-minute summed images of the dynamic scan and for the torso survey. For each lesion, region-of-interest mean and maximum SUV and SUL were recorded, with LBM estimated from height (cm) and weight (kg; ref. 17). Results of a "draw–redraw" reader study in a subset of scans showed excellent agreement for SUV quantitation for both FES and FDG PET (18).

Study treatment and monitoring of response

Clinical follow-up and response assessment followed standard of care at UW/SCCA and included additional clinical FDG PET scans, conventional imaging (e.g., CT, MRI, bone scan), serum tumor markers, and evaluation of symptoms, as deemed appropriate by the treating physician. PFS was measured from the date of the FES PET scan to the date of progressive disease or death from any cause. For patients who did not experience progressive disease or were lost to follow-up more than 4 months before death, PFS was censored at the date of the last radiographic assessment that did not show progressive disease or at surgery for patients receiving neoadjuvant endocrine therapy. OS was measured from the date of FES PET scan until death and was censored at the last date the patient was known to be alive. Clinical benefit was defined as 6 or more months on endocrine therapy without progressive disease or death (PFS6) and was missing if a patient had <6 months of follow-up.

Statistical analysis

Imaging biomarkers evaluated as predictors of PFS, OS, and 6-month response. A primary aim of this study was to validate FES SUV $\leq$1.5 as a classifier to identify patients who would be unlikely to benefit from endocrine therapy. The cut-off point of FES SUV $\leq$1.5 (for the average uptake for up to three lesions in the dynamic scan FOV, "FES dynSUV") had been reported by our group as a threshold under which patients were unlikely to benefit from endocrine therapy (11). Although the average uptake for up to three lesions is likely a good patient level summary (18), subsequent development of FES PET measures has suggested alternative parameterizations to optimize performance of FES quantitative biomarkers. These alternatives include the use of SUVs measured on the torso survey (19, 20), LBM adjustments (15), and log transformation prior to averaging to limit the influence of extreme measurements (21).

Another proposed alternative parameterization is the ratio of FES SUV_{max} and FDG SUV_{max}. Weighting the FES uptake by metabolic aggressiveness (as measured by FDG SUV) could serve as an "index of differentiation" (22). In addition, as both FDG and FES uptake measures are similarly impacted by partial-volume effects, the ratio of FES to FDG uptake measures minimizes tumor partial volume effects. Therefore, we prospectively defined additional patient-level measures to test for associations with PFS, OS, and PFS6:

(1) Geometric mean of FES SUL_{mean} (LBM-adjusted SUV) for the 3 lesions with the highest FDG SUV_{max} in the torso scan (FES SUL_{mean3}, cut-off point 1.0): For the torso scan portion of the FES PET sequence, an SUL_{mean} of 1.0 or 1.1 was found to be roughly equivalent to the dynamic scan SUV_{mean} of 1.5 (18). The geometric mean was used to accommodate log transformation. The average value for 3 lesions was selected because within-patient FES uptake did not vary greatly by lesion (18), and assessment of a large number of lesions does not improve quantitation in settings such as the RECIST criteria for measuring treatment response (23). The three lesions with highest FDG uptake were selected as the most relevant for therapy selection, as targeting of an indolent FES-avid lesion would not be successful therapy if an aggressive, resistant lesion remained.

(2) Presence/absence of lesion(s) with FES $SUL_{mean} \leq 1$ and FDG $SUV_{max} > 5$: The classification of the presence/absence of aggressive lesions was based on the threshold for low FES SUL_{mean} (11, 18) and an arbitrary threshold of 5 for a "high" FDG SUV_{max}.

(3) Presence/absence of tumor site(s) qualitatively negative for FES uptake: The qualitative assessment of FES uptake was

assessed as a straightforward measure that would take less time to assess than quantitative measures. Again, the focus was on identifying FES-negative lesions hypothesized to be resistant to endocrine therapy.

Analysis plan. The primary analysis evaluated the incremental value of summary imaging measures for predicting PFS. First, Cox proportional hazards regression was used to evaluate patient characteristics (age), disease characteristics (months since primary cancer diagnosis, number of tumors in the torso scan, number of prior treatment regimens for metastatic disease, presence/absence of visceral disease), and tumor characteristics (PgR, histology) as potential predictors. Then, each imaging biomarker was assessed for additional contributions. Secondary response measures were OS and clinical benefit (PFS6). Predictors were evaluated using Cox proportional hazards regression and logistic regression, with all tests two-sided.

Exploratory analyses used tree-based methods to identify clinically important subgroups potentially identified by combinations of factors (24). Patient characteristics, disease history, tumor histology, and imaging biomarkers were assessed as predictors of PFS using recursive partitioning, choosing splits based on reduction of a one-step deviance (25) and pruning to minimize cross-validation prediction error (26). Statistical analyses were conducted using SAS/STAT software, version 9.4 (SAS Institute, Inc.), including the %ICC9 macro (for concordance analysis reported in the Supplementary Data; ref. 27) and R version 3.1.3 (R Foundation for Statistical Computing), including the "rpart package".

Results

FES PET scans were conducted in 110 patients scheduled for initiation of or change in endocrine therapy, between January 2000 and January 2011. Ninety-one patients' scans were from an earlier series evaluating between- and within-patient FES uptake heterogeneity (18), and 19 were previously described in a single-center IND study (16). Twenty patients were excluded from this analysis. Thirteen had $HER2^+$ disease. Four others received additional or alternate therapy, 2 had adverse events (pneumonia, surgical complications) soon after imaging, and one did not have dynamic FES PET data. Approximately 35 of the remaining patients appeared in a prior analysis (11), but that study examined objective response rather than PFS, OS, or PFS6.

Patient and disease characteristics are described in Table 1 for the 90 patients eligible for response analysis. The mean age at the time of FES PET was 56 years (SD, 11 years), and patients ranged in age from 28 to 79 years old. Time from breast cancer diagnosis to FES PET ranged from 0 months to more than 20 years, with a median of 4 years and 0 to 5 prior chemotherapy regimens for metastatic disease. The number of tumors identified by FES PET and FDG PET was 1 to 16, with a median of 5 tumors per patient. Sixty-seven patients had a date of progression identified, with no deaths as PFS events. The 57 deaths recorded included 11 patients whose PFS was censored (at 0–7 months).

Table 1. Patient and disease characteristics as recorded at time of FES PET scan and clinical outcomes ($N = 90$)

	***N* (%)**
Age	
<50 years	26 (29%)
≥50 years	64 (71%)
Sex	
Female	88 (98%)
Male	2 (2%)
Menopausal status	
Premenopausal	14 (16%)
Postmenopausal/male	76 (84%)
Time since breast cancer diagnosis	
<2 months	22 (24%)
2 months–5 years	31 (34%)
>5 years	37 (41%)
Primary tumor PgR	
Positive or weakly positive	72 (80%)
Negative	14 (16%)
Not assessed	4 (4%)
Histology of primary breast cancer	
Ductal	66 (73%)
Lobular	17 (19%)
Ductal and lobular	6 (7%)
Unknown	1 (1%)
Visceral disease at time of scan	
Liver, lung, stomach lesion(s)	14 (16%)
No liver, lung, stomach lesion(s)	76 (84%)
Endocrine therapy following FES PET	
Aromatase inhibitor	56 (62%)
Aromatase inhibitor + fulvestrant	20 (22%)
Tamoxifen	9 (10%)
Ovarian suppression only	2 (2%)
Fulvestrant	2 (2%)
Unknown	1 (1%)
PFS	
Events	67 (74%, range 1–50 months)
Censored	23 (26%, range 0–54 months)
OS	
Events (deaths)	57 (63%, range 5–128 months)
Censored	33 (37%, range 3–130 months)
Clinical benefit	
PFS > 6 months	47 (52%, 60% of assessed)
PFS ≤ 6 months	31 (34%)
Not assessed[a]	12 (13%)
	Mean (SD), range
FES dynSUV	2.73 (1.80), 0.50–9.62
FES SUL_{mean3}[b]	1.80 (1.35), 0.17–6.67
FES SUV_{max3}[b]	2.84 (1.79), 0.57–8.79
FDG SUL_{max3}[b]	3.69 (2.43), 0.77–15.16
FDG SUV_{max3}[b]	5.78 (3.74), 1.25–22.95
FES/FDG_{ratio3}[c]	0.92 (0.29), 0.34–1.71

NOTE: FES dynSUV average of FES SUV_{mean} of up to 3 lesions in the dynamic FOV.

[a]Seven patients had surgery or chemotherapy within 6 months of FES PET scan and without disease progression; 5 others were lost to follow-up within 6 months of FES PET scan.

[b]FES $SUL_{mean3} = 2\left[\frac{\sum_1^{n_i} log_2(\text{FES SULmean})}{n_i}\right]$, geometric mean for up to 3 lesions in torso sweep with highest FDG SUV_{max}. Others are similarly constructed geometric means.

[c]FES/FDG ratio3 $= \sum_1^{n_i}\left(\sqrt{\frac{\text{FES SUVmax}}{\text{FDG SUVmax}}}\right)/n_i$, average of square root of measure for up to 3 lesions in torso sweep with highest FDG SUV_{max}.

Imaging classifiers of endocrine resistance

Table 2 shows the distributions for four proposed classifiers of endocrine resistance. Measures based on the quantitative FES and FDG PET had 20% to 29% rate of predicted endocrine resistance. The rate was lower for the qualitative reads, with lesions lacking FES uptake above background detected in only 12 of 90 scans (13%; 95% confidence interval, 8%–22%). More than half of patients (51/90) had tumors with sufficiently high FES uptake for

Table 2. Descriptive statistics for predefined classifiers for endocrine resistance ($N = 90$)

Classifier	Description	*N* (%)
FES dynSUV $\leq$ 1.5	FES dynSUV = average of FES SUV_{mean} for up to 3 lesions in the dynamic FOV	24/90 (27%)
FES $SUL_{mean3} \leq 1$	See Table 1[a]	26/90 (29%)
Any low FES/FDG	Any low FES/FDG = presence of any lesion with FES $SUL_{mean} \leq 1$ and FDG $SUV_{max} > 5$	17/88[a] (19%)
Qualitative FES neg	Qualitative FES negative = presence of any lesion lacking FES uptake above background (qualitative read)	12/90 (13%)

[a]Two patients had FDG PET scan from outside sites, used to locate lesions but not for uptake quantification; to calculate SUL_{mean3} for these patients, 3 lesions were selected at random (from 5 and 8 lesions).

all the four classifiers to indicate likely response to endocrine therapy; in contrast, classification of likely nonresponders among the remaining patients varied greatly. Detailed measures of agreement are reported in Supplementary Data.

Prediction of response to endocrine therapy

PFS was assessed in 84 patients, excluding 2 patients without quantitative FDG PET and 4 patients followed only for OS. Analysis of PFS6 excluded an additional 8 patients censored for PFS before 6 months.

Before evaluating PET classifiers to predict PFS, we examined patient and disease characteristics. Using statistical significance (likelihood ratio test $P < 0.05$) as a criterion, age, presence of visceral disease, number of prior chemotherapy regimens for metastatic disease, and tumor histology were dismissed as predictors of PFS in this cohort. Additional predictors associated with OS but not PFS in univariate models were PgR (HR, 3.3 for negative; $P = 0.002$), burden of disease as measured by log-transformed number of lesions (HR, 1.8; $P = 0.002$), and log-transformed number of months from cancer diagnosis to FES PET scan (HR, 1.2; $P = 0.01$). All remained statistically significant in a multivariable model predicting OS.

None of the four preselected binary PET classifiers predicted PFS (Fig. 1, log-rank $P > 0.05$). They also did not contribute beyond PgR, time since diagnosis, and number of lesions for predicting OS (likelihood-ratio $P > 0.10$). Fully quantitative individual measures (without prespecified cut-off points) are displayed in Supplementary Fig. S2 (Supplementary Data) and do not show a strong association with clinical benefit (PFS6).

Because the prospectively planned analysis did not yield a clear answer for further development of biomarkers for endocrine therapy response, we conducted an exploratory analysis using data-driven criteria to select the "best" predictors of PFS. Predictors examined included patient (age) and clinical variables (PgR, number of lesions, visceral disease, number of prior chemotherapy regimens for metastatic disease), and imaging parameters (FES dynSUV, FES SUL_{mean3}, FES/FDG ratio3, qualitative FES). Additional imaging characteristics were included, each evaluated as a geometric mean for up to 3 lesions in the torso sweep with highest FDG SUV_{max} (FES SUL_{max3}, FES SUV_{max3}, FDG SUL_{max3}, FDG SUV_{max3}).

Recursive partitioning with equal costs for all covariates, selecting the number of splits that minimized the 5-fold

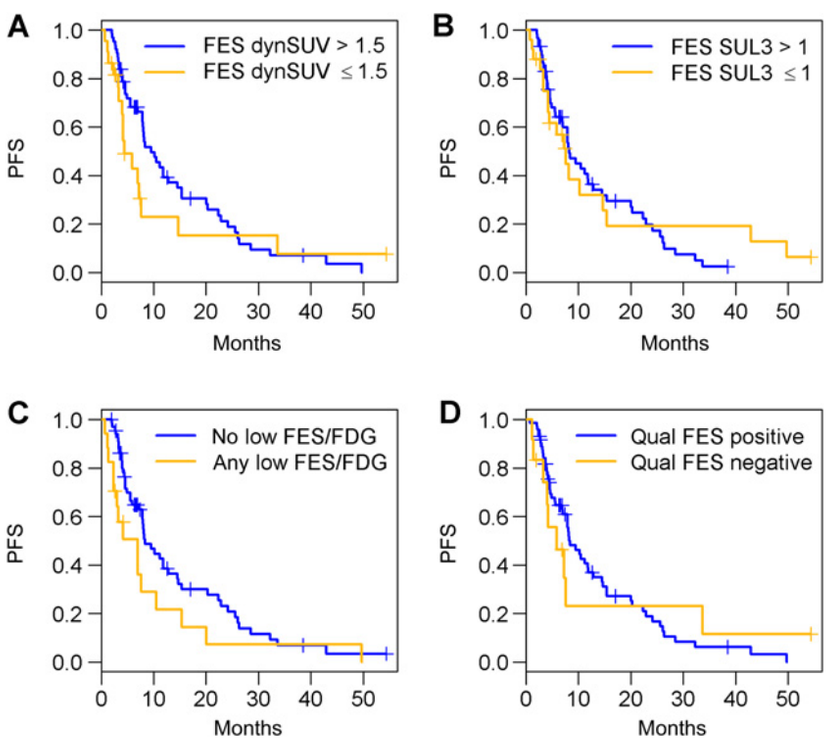

Figure 1.
Kaplan–Meier curves of PFS for prespecified quantitative imaging classifiers predicting endocrine response ($N = 84$, $P > 0.05$ for all). **A,** FES dynSUV = average of FES SUV_{mean} of up to 3 lesions in the dynamic FOV. **B,** FES SUL3 = geometric mean of FES SUL_{mean} (LBM-adjusted FES SUV_{mean}) for up to 3 lesions in torso sweep with highest FDG SUV_{max}. **C,** Any low FES/FDG = presence of any lesion with FES $SUL_{mean} \leq 1$ and FDG $SUV_{max} > 5$. **D,** Qualitative (Qual) FES negative = presence of any lesion lacking FES uptake above background (qualitative read).

Table 3. Prediction of PFS, OS, and clinical benefit (PFS > 6 months on endocrine therapy) by a 3-level classifier based on FDG and FES PET results ("group")

Group	*N* (PFS)	Median months PFS (95% confidence interval)	*N* (OS)	Median months OS (95% confidence interval)	PFS $\geq$ 6 months (95% confidence interval)
FDG SUL_{max3} < 2.2	24	26.1 (11.2–49.7)	24	69.5 (55.1–NE)	18/21 = 86% (65%–95%)
FDG $SUL_{max3} \geq$ 2.2 FES $SUL_{mean3} \geq$ 0.85	50	7.9 (5.6–11.8)	53	41.9 (36.3–64.7)	27/46 = 59% (44%–72%)
FDG $SUL_{max3} \geq$ 2.2 FES SUL_{mean3} < 0.85	10	3.3 (1.4–NE)	11	36.8 (22.1–NE)	1/9 = 11% (1%–43%)

NOTE: This classifier was selected by recursive partitioning with 5-fold internal cross-validation as the best predictor of PFS among patient (age) and clinical variables (PgR, number of lesions, visceral disease, number of prior chemotherapy regimens for metastatic disease) and imaging parameters (FES SUL_{mean3}, FES/FDG ratio, qualitative FES, and alternate summaries such as FES SUV_{max3}).
N = 88 for OS excludes 2 without FDG SUL_{max} quantified.
N = 84 for PFS also excludes 4 without follow-up for progression.
N = 76 for clinical benefit also excludes 8 censored for progression before 6 months.
FDG SUL_{max3} = geometric mean of FDG SUL_{max} for 3 lesions in torso sweep with highest FDG SUV_{max}.
FES SUL_{mean} = geometric mean of FES SUL_{mean} for 3 lesions in torso sweep with highest FDG SUV_{max} (see Table 1 footnotes).
Abbreviation: NE, not evaluable.

cross-validated prediction error (26), led to a model with FDG SUL_{max3} and FES dynSUV predicting PFS. Patients with low FDG uptake suggesting indolent tumors (FDG SUL_{max3} < 2.2) had the most favorable PFS; among patients whose lesions were more FDG-avid, patients with FES-avid lesions had more favorable PFS on endocrine therapy. The FES dynSUV cut-off point of 1.47 was essentially the same as previously identified for predicting objective response (11). However, as dynamic FES PET protocols are not considered essential for quantification of FES uptake (15) and 30-minute summed images are not clinically practical for a single FOV, recursive partitioning was repeated without FES dynSUV, selecting FDG SUL_{max3} and FES SUL_{mean3} (Table 3). Evaluated by Kaplan–Meier methods and the log-rank test, the three groups identified by recursive partitioning had markedly different PFS (Fig. 2A, $P < 0.001$). These groups also predicted OS (Fig. 2B, $P = 0.01$) and PFS6 (Table 3, Wald test, $P < 0.001$).

Twenty-four of the patients (27%) were identified by FDG PET as having indolent disease, with a median PFS of 26.1 months, 86% rate of clinical benefit (≥6 months of response or stable disease on endocrine therapy), and median OS of more than 5 years. A representative case is displayed in Fig. 3A. FES PET was used to stratify the remaining patients' expected benefit from endocrine therapy. Representative cases are displayed for patients with high FDG SUL_{max3} and high FES SUL_{mean3} (Fig. 3B) and high FDG SUL_{max3} and low FES SUL_{mean3} (Fig. 3C). The dramatic difference in median PFS detected through tree-based exploratory analysis (7.9 vs. 3.3 months for high versus low FES SUL_{mean3}) appeared to apply to clinical benefit (59% vs. 11%), but the groups did not differ in median OS (42 vs. 37 months; Table 3; Fig. 2). When added to a Cox regression model predicting OS, the 3-way PET classifier was an independent predictor beyond PgR, time since diagnosis, and number of lesions for predicting OS (likelihood-ratio $P = 0.046$). Exclusion of the 5 patients with early-stage disease did not affect the cut-off points for the 3-way classifier for PFS; additional sensitivity analyses applying recursive partitioning to OS and PFS6 endpoints, and models only considering SUV_{max} and SUL_{max} imaging parameters, are described in the Supplementary Data.

Discussion

We analyzed imaging characteristics and tumor response in ER^+ breast cancer patients undergoing endocrine therapy in a mixed first-line and salvage setting. We prespecified 4 potential imaging biomarker classifiers to predict response to endocrine therapy. The classifiers used information from FDG PET and/or

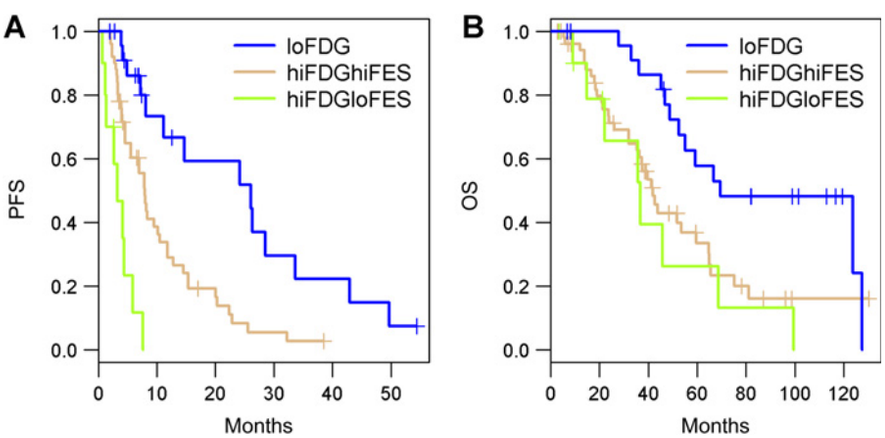

Figure 2.
Results of recursive partitioning to classify endocrine response (see Table 3). Kaplan–Meier curves. **A,** PFS predicted by FES SUL_{mean3} and FDG SUL_{max3} (log-rank test $P < 0.001$). **B,** OS predicted by FES SUL_{mean3} and FDG SUL_{max3} (log-rank test $P = 0.01$).

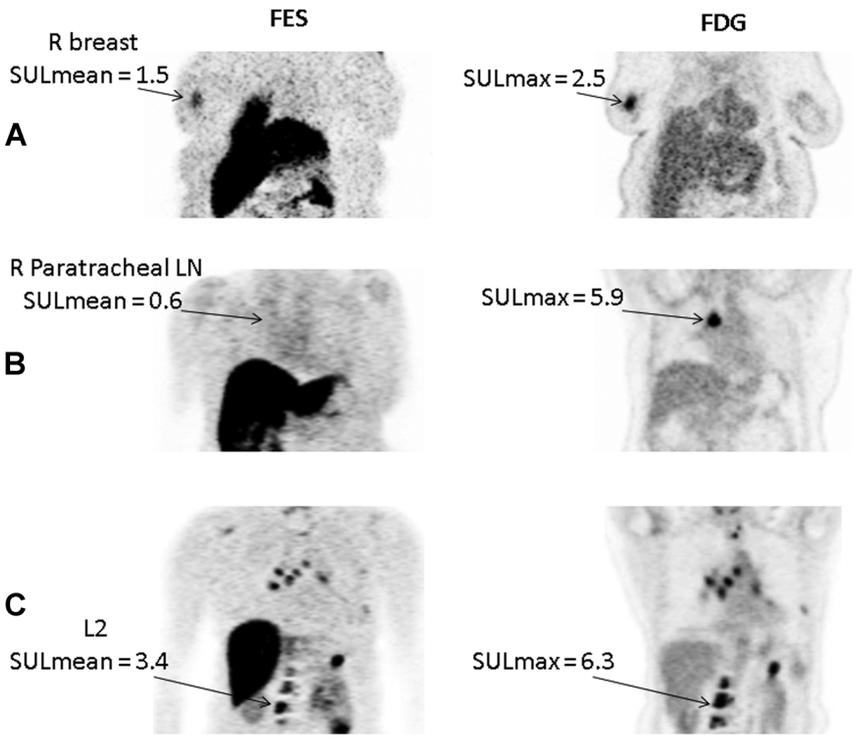

Figure 3.

Representative cases for the three classification groups in Table 3 and Fig. 2. Coronal view from FES PET and FDG PET scans. Ordered subset expectation maximization (OSEM) reconstruction was used for improved presentation, but not for quantitation. **A,** Low FDG SUL_{max3}. This 56-year-old woman had 4 lesions in her breast and lymph nodes (LN). Geometric mean FDG SUL_{max} for the 3 hottest lesions was 1.5 (geometric mean FES SUL_{mean} was 1.1). She was on anastrozole for 5 months until progression. R, Right. **B,** High FDG SUL_{max3} and low FES SUL_{mean3}. This 59-year-old woman had 5 lesions in lymph nodes and spine. Geometric mean FDG SUL_{max} for the 3 hottest lesions was 4.4. Geometric mean FES SUL_{mean3} was 0.3. She was on tamoxifen for 3 months until progression. **C,** High FDG SUL_{max3} and high FES SUL_{mean3}. This 59-year-old woman had lesions in the breast, chest wall, and hilar nodes, as well as multiple bony lesions. Geometric mean FDG SUL_{max} for the 3 hottest lesions was 12.7. Geometric mean FES SUL_{mean} was 6.6. She was on exemestane for 9.5 months until progression.

FES PET and included a qualitative FES PET measure. Although none of these classifiers predicted PFS or OS in multivariate models controlling for known risk factors, exploratory tree-based methods revealed a novel 3-way classification using both FDG and FES PET results. In this heterogeneous and heterogeneously treated population, with varying disease stage and lines of prior therapy, about 13% (11/88) of patients displayed a phenotype (high average FDG SUL_{max3}, low average FES SUL_{mean3}) that was unlikely to benefit from endocrine therapy. Our analysis suggested that FES and FDG PET imaging can help guide therapy selection, or therapy dosing (28), and assist clinicians who face an expanding array of therapeutic options. Sorting patients into 3 groups (indolent disease, aggressive but strongly ER avid, and aggressive and weakly ER avid) has relevance for clinicians and patients weighing toxicities, costs, and logistics of therapeutic options. For instance, an FES PET scan would not be necessary to predict outcome if the FDG PET indicates indolent disease.

Our results suggest that for ER^+ (primary) breast cancer, patients with indolent disease by FDG PET are potential candidates for endocrine therapy, as are patients with more aggressive disease but high ER activity as detected by FES PET. Although the 3-way classifier is novel, our results are consistent with and complementary to prior studies predicting response to endocrine therapy. Average FES SUV_{mean} of about 1.5 from a single-FOV dynamic scan, the same classifier used in earlier studies to predict clinical response (11), predicted superior PFS but only in patients with nonindolent (FDG $SUL_{max3} \geq 2.2$) disease. A nearly identical cut-off point for FDG SUL_{max} was selected to predict OS, even though none of the PFS events were deaths. FES cut-off points of 0.85 (SUL_{mean3}) and 1.5 (dynSUV) may seem too close to background activity to allow reproducible quantification, but background activity for FES is <1 due to intense uptake in the liver. Some expected predictors of response to endocrine therapy, such as primary tumor PgR expression, did not predict PFS or OS in this heterogeneous, moderately sized cohort.

This study has several limitations. First, patient disease characteristics in this heterogeneous patient population arising from four different studies varied greatly beyond having had ER^+ primary breast cancer. At the time, these patients were all candidates for endocrine therapy, but these studies pre-dated modern molecularly targeted drugs. We did not take into account prior benefit from endocrine therapy or indolent disease as suggested by late recurrence. There also may be period effects in supportive care, treatment options, and patient selection due to the 12-year range of recruitment. The use of older PET reconstruction methods was necessary for attempted validation of prior quantitative uptake measure cut-off points but may limit the generalizability of our classification. Although the sample size is relatively large (and outcomes data relatively complete) for a study of novel biomarkers that cannot be assessed retrospectively, it is too small for rigorous internal validation of cut-off points, and external validation is required as well. Our analysis also does not consider tumor genomic profiles or account for acquired mutations in ER (29–31), as these assays were not performed routinely during that time. However, this study reflects key advantages of imaging. First, imaging assesses the entire burden of disease rather than a sample of tissue or plasma tumor DNA. In addition, patients may refuse biopsy due to fears of pain or treatment delay, sampling may not be feasible, and tissue sampled from metastatic disease is not analyzable in many cases (32). An imaging biomarker trial applying FES PET and FDG PET prospectively in a multicenter study (NCT02398773) will further evaluate qualitative and quantitative PET parameters as predictors of endocrine therapy response.

In conclusion, our data suggest that the combination of FDG PET and FES PET will be useful in discriminating between patients with disease that is indolent or amenable to endocrine therapy and patients whose disease is more aggressive and is unlikely to respond to endocrine therapy. Further confirmation with a prospective study design and modern options for endocrine and molecularly targeted treatment is needed to test the ability of combined FDG and FES PET to guide therapy selection for patients with ER^+ metastatic breast cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: B.F. Kurland, L.M. Peterson, E.K. Schubert, K.A. Krohn, D.A. Mankoff, H.M. Linden

Development of methodology: L.M. Peterson, D.A. Mankoff, H.M. Linden

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): L.M. Peterson, E.K. Schubert, J.M. Link, D.A. Mankoff, H.M. Linden

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): B.F. Kurland, L.M. Peterson, J.H. Lee, J.M. Link, D.A. Mankoff

Writing, review, and/or revision of the manuscript: B.F. Kurland, L.M. Peterson, J.H. Lee, E.R. Currin, J.M. Link, D.A. Mankoff, H.M. Linden

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): B.F. Kurland, L.M. Peterson, E.R. Currin

Acknowledgments

We thank Thomas Lawton, Kim Allison, Mark Muzi, Paul Kinahan, the study participants, the referring physicians, and the nuclear medicine team in the University of Washington Radiology Department.

Grant Support

This study was supported by a Fred Hutchinson Cancer Research Center Breast Cancer Research Program Pilot grant (Safeway), NCI Phase I and II Cancer Imaging Program (contract #N01-CM-37008, Protocol 8052), Komen SAC140060, NIH grants CA42045 and CA148131, and DOD W81XWH-04-01-0675. This project used the UPCI Biostatistics Shared Resource Facility that is supported in part by award P30CA047904.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 8, 2016; revised May 5, 2016; accepted June 7, 2016; published OnlineFirst June 24, 2016.

References

1. Kwast AB, Voogd AC, Menke-Pluijmers MB, Linn SC, Sonke GS, Kiemeney LA, et al. Prognostic factors for survival in metastatic breast cancer by hormone receptor status. Breast Cancer Res Treat 2014;145:503–11.
2. Baselga J, Campone M, Piccart M, Burris HAIII, Rugo HS, Sahmoud T, et al. Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer. N Engl J Med 2012;366:520–9.
3. Turner NC, Ro J, Andre F, Loi S, Verma S, Iwata H, et al. Palbociclib in hormone-receptor-positive advanced breast cancer. N Engl J Med 2015;373:209–19.
4. Johnston SR. BOLERO-2 - will this change practice in advanced breast cancer? Breast Cancer Res 2012;14:311.
5. Early Breast Cancer Trialists' Collaborative Group, Davies C, Godwin J, Gray R, Clarke M, Cutter D, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011;378:771–84.
6. Mintun MA, Welch MJ, Siegel BA, Mathias CJ, Brodack JW, McGuire AH, et al. Breast cancer: PET imaging of estrogen receptors. Radiology 1988;169:45–8.
7. Peterson LM, Mankoff DA, Lawton T, Yagle K, Schubert EK, Stekhova S, et al. Quantitative imaging of estrogen receptor expression in breast cancer with PET and 18F-fluoroestradiol. J Nucl Med 2008;49:367–74.
8. Gemignani ML, Patil S, Seshan VE, Sampson M, Humm JL, Lewis JS, et al. Feasibility and predictability of perioperative PET and estrogen receptor ligand in patients with invasive breast cancer. J Nucl Med 2013;54:1697–702.
9. Mortimer JE, Dehdashti F, Siegel BA, Katzenellenbogen JA, Fracasso P, Welch MJ. Positron emission tomography with 2-[18F]Fluoro-2-deoxy-D-glucose and 16alpha-[18F]fluoro-17beta-estradiol in breast cancer: correlation with estrogen receptor status and response to systemic therapy. Clin Cancer Res 1996;2:933–9.
10. Mortimer JE, Dehdashti F, Siegel BA, Trinkaus K, Katzenellenbogen JA, Welch MJ. Metabolic flare: indicator of hormone responsiveness in advanced breast cancer. J Clin Oncol 2001;19:2797–803.
11. Linden HM, Stekhova SA, Link JM, Gralow JR, Livingston RB, Ellis GK, et al. Quantitative fluoroestradiol positron emission tomography imaging predicts response to endocrine treatment in breast cancer. J Clin Oncol 2006;24:2793–9.
12. Ulaner GA, Eaton A, Morris PG, Lilienstein J, Jhaveri K, Patil S, et al. Prognostic value of quantitative fluorodeoxyglucose measurements in newly diagnosed metastatic breast cancer. Cancer Med 2013;2:725–33.
13. Daube-Witherspoon ME, Karp JS, Casey ME, DiFilippo FP, Hines H, Muehllehner G, et al. PET performance measurements using the NEMA NU 2-2001 standard. J Nucl Med 2002;43:1398–409.
14. Hamacher K, Coenen HH, Stocklin G. Efficient stereospecific synthesis of no-carrier-added 2-[18F]-fluoro-2-deoxy-D-glucose using aminopolyether supported nucleophilic substitution. J Nucl Med 1986;27:235–8.
15. Peterson LM, Kurland BF, Link JM, Schubert EK, Stekhova S, Linden HM, et al. Factors influencing the uptake of 18F-fluoroestradiol in patients with estrogen receptor positive breast cancer. Nucl Med Biol 2011;38:969–78.
16. Peterson LM, Kurland BF, Schubert EK, Link JM, Gadi VK, Specht JM, et al. A phase 2 study of 16alpha-[18F]-fluoro-17beta-estradiol positron emission tomography (FES-PET) as a marker of hormone sensitivity in metastatic breast cancer (MBC). Mol Imaging Biol 2014;16:431–40.
17. Hallynck TH, Soep HH, Thomis JA, Boelaert J, Daneels R, Dettli L. Should clearance be normalised to body surface or to lean body mass? Br J Clin Pharmacol 1981;11:523–6.
18. Kurland BF, Peterson LM, Lee JH, Linden HM, Schubert EK, Dunnwald LK, et al. Between-patient and within-patient (site-to-site) variability in estrogen receptor binding, measured *in vivo* by 18F-fluoroestradiol PET. J Nucl Med 2011;52:1541–9.
19. Dehdashti F, Mortimer JE, Trinkaus K, Naughton MJ, Ellis M, Katzenellenbogen JA, et al. PET-based estradiol challenge as a predictive biomarker of response to endocrine therapy in women with estrogen-receptor-positive breast cancer. Breast Cancer Res Treat 2009;113:509–17.
20. van Kruchten M, Glaudemans AW, de Vries EF, Beets-Tan RG, Schroder CP, Dierckx RA, et al. PET imaging of estrogen receptors as a diagnostic tool for breast cancer patients presenting with a clinical dilemma. J Nucl Med 2012;53:182–90.
21. Thie JA, Hubner KF, Smith GT. The diagnostic utility of the lognormal behavior of PET standardized uptake values in tumors. J Nucl Med 2000;41:1664–72.
22. Mankoff DA, Dehdashti F. Imaging tumor phenotype: 1 plus 1 is more than 2. J Nucl Med 2009;50:1567–9.
23. Hillman SL, An MW, O'Connell MJ, Goldberg RM, Schaefer P, Buckner JC, et al. Evaluation of the optimal number of lesions needed for tumor evaluation using the Response Evaluation Criteria in Solid Tumors: a North Central Cancer Treatment Group investigation. J Clin Oncol 2009;27:3205–10.
24. Clark GM, Hilsenbeck SG, Ravdin PM, De Laurentiis M, Osborne CK. Prognostic factors: rationale and methods of analysis and integration. Breast Cancer Res Treat 1994;32:105–12.

25. LeBlanc M, Crowley J. Relative risk trees for censored survival data. Biometrics 1992;48:411–25.
26. Everitt B, Hothorn T. A handbook of statistical analyses using R. Boca Raton, FL: CRC Press; 2006.
27. Hankinson SE, Manson JE, Spiegelman D, Willett WC, Longcope C, Speizer FE. Reproducibility of plasma hormone levels in postmenopausal women over a 2–3-year period. Cancer Epidemiol Biomarkers Prev 1995;4:649–54.
28. Heidari P, Deng F, Esfahani SA, Leece AK, Shoup TM, Vasdev N, et al. Pharmacodynamic imaging guides dosing of a selective estrogen receptor degrader. Clin Cancer Res 2015;21:1340–7.
29. Schiavon G, Hrebien S, Garcia-Murillas I, Cutts RJ, Pearson A, Tarazona N, et al. Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. Sci Transl Med 2015;7:313ra182.
30. Wang P, Bahreini A, Gyanchandani R, Lucas PC, Hartmaier RJ, Watters RJ, et al. Sensitive detection of mono- and polyclonal ESR1 mutations in primary tumors, metastatic lesions, and cell-free DNA of breast cancer patients. Clin Cancer Res 2016;22:1130–7.
31. Leyland-Jones B, Gray KP, Abramovitz M, Bouzyk M, Young B, Long B, et al. ESR1 and ESR2 polymorphisms in the BIG 1-98 trial comparing adjuvant letrozole versus tamoxifen or their sequence for early breast cancer. Breast Cancer Res Treat 2015;154:543–55.
32. Amir E, Miller N, Geddie W, Freedman O, Kassam F, Simmons C, et al. Prospective study evaluating the impact of tissue confirmation of metastatic disease in patients with breast cancer. J Clin Oncol 2012;30:587–92.