Journal of Nuclear Medicine, published on September 10, 2015 as doi:10.2967/jnumed.115.160663

A Phase II Study Of [(18)F]-3'deoxy-3'-Fluorothymidine Positron Emission Tomography (FLT-PET) In The Assessment Of Early Response Of Breast Cancer To Neoadjuvant Chemotherapy: Results From ACRIN 6688

Lale Kostakoglu[1], Fenghai Duan[2], Michael O. Idowu[3], Paul R. Jolles[3], Harry D. Bear[3,4], Mark Muzi[5], Jean Cormack[2], John P. Muzi[5], Daniel A. Pryma[6], Jennifer M. Specht[5], Linda Hovanessian-Larsen[7], John Miliziano[8], Sharon Mallett[9], Anthony F. Shields[10], David A. Mankoff[6]

[1]Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, NY;
[2]Department of Biostatistics and Center for Statistical Sciences, Brown University School of Public Health, Providence, RI;
[3]Virginia Commonwealth University, Richmond, VA;
[4]Massey Cancer Center of Virginia Commonwealth University, Richmond, VA;
[5]University of Washington, Seattle, WA;
[6]Abramson Cancer Center and Perelman School of Medicine University of Pennsylvania, Philadelphia, PA;
[7]University of Southern California, Los Angeles, CA;
[8]Morton Mease Plant Hospital, Clearwater, FL;
[9]American College of Radiology Imaging network (ACRIN), Philadelphia, PA;
[10]Karmanos Cancer Institute, Wayne State University, Detroit, MI;

This work was supported by ACRIN, which receives funding from the National Cancer Institute through U01 CA080098, under the American Recovery and Reinvestment Act of 2009 (ARRA) and U01 CA079778.

Running title: FLT-PET in early response of breast cancer

Key words: Early treatment response, ^{18}F-FLT-PET, breast cancer, neoadjuvant therapy

Abstract

Our objective was to determine whether early change in standardized uptake values (SUV) of [(18)F]-3'deoxy-3'-fluorothymidine using positron emission tomography with computed tomography (^{18}F-FLT-PET/CT) can predict pathologic complete response (pCR) of primary breast cancer to neoadjuvant chemotherapy (NAC). The key secondary objective was to correlate SUV with the proliferation marker, Ki-67, at baseline and after NAC. Methods: This prospective, multicenter phase II study did not specify the therapeutic regimen, thus, NAC varied among centers. All evaluable patients underwent ^{18}F-FLT - PET/CT at baseline (FLT1), and after one cycle of NAC (FLT2); 43 patients were imaged at FLT1, FLT2 and after NAC completion (FLT3). The percentage change in SUV_{max} ($\%\Delta SUV_{max}$) between FLT1 and FLT2 and FLT3 was calculated for the primary tumors. The predictive value of ΔSUV_{max} for pCR was determined using the receiver operator curve (ROC) analysis. The correlation between SUV_{max} and Ki-67 was also assessed. Results: A total of 51 of 90 recruited patients [median age 54; stage IIA-IIIC] met the eligibility criteria for the primary objective analysis with an additional 22 patients totaling 73 patients for secondary analyses. A pCR in the primary breast cancer was achieved in 9 of 51 pts. NAC resulted in a significant reduction in $\%SUV_{max}$ (mean Δ, 39%; 95%CI 31–46). There was a marginal difference in $\%\Delta SUV_{max\_FLT1\text{-}FLT2}$ between pCR and no-pCR patient groups (Wilcoxon one-sided p=0.050). The AUC for ΔSUV_{max} in the prediction of pCR was 0.68 [90%CI 0.50 - 0.83; Delong's one-sided p=0.05], with slightly better predictive value for $\%SUV_{mean}$ (p=0.02) and similar prediction for SUV_{peak} (p=0.04). There was a weak correlation with pre-therapy SUV_{max} and Ki-67 (r=0.29, p=0.04), but the correlation between SUV_{max} and Ki-67 after completion of NAC was stronger (r=0.68, P< 0.0001). Conclusions: FLT-PET imaging of breast to test ^{18}F-FLT-PET/CT in a more uniformly treated patient population.

INTRODUCTION

Systemic neoadjuvant chemotherapy (NAC) before surgery plays a role in locally advanced breast cancer (LABC) to downstage disease to increase the chances for breast-conserving surgery, eradicate micrometastases, and provide an indication of therapeutic responsiveness (1). The eradication of invasive cancer after NAC, i.e. pathologic complete response (pCR), is a predictor of improved survival (2, 3). However, conventional chemotherapy regimens result in pCR in only a minority of patients. In the era of individualized medicine, assessing pathologic response after completion of NAC does not allow for therapy adaptation with earlier discontinuation of ineffective therapies. Patients without sufficient response might benefit from switching to more effective alternative treatments early during the course of NAC. In this regard, there is a need to develop non-invasive imaging methodologies that can provide an early indication of response. Currently, treatment response is largely assessed by measurement of tumor size after several cycles of chemotherapy. Size changes can lag behind therapy-induced molecular changes, motivating the use of molecular imaging methods to assess response. Although, 18F-fluorodeoxyglucose (FDG) PET has been used method for monitoring response to treatment in breast cancer, it has limitations that include prediction of pCR across different phenotypes (4, 5). PET imaging with 3'-deoxy-3'-18F-fluorothymidine (^{18}F-FLT-PET) provides a noninvasive method for evaluating cell proliferation, an early indicator of therapeutic response. FLT is a substrate for thymidine kinase-1 and the accumulation of FLT in tumors provides a quantitative measure of cell proliferation through the relationship between thymidine kinase-1 expression and cell cycle regulation (6-9). Several prior single-center studies have demonstrated that changes in breast cancer tumor proliferation assessed by ^{18}F-FLT--PET/CT early after initiating chemotherapy predict tumor response with good sensitivity, however, the results of these studies were variable and none of the prior studies investigated the potential for predicting pCR to NAC (10, 11-17). In the present multicenter study, our objective was to correlate changes measured by FLT in the primary tumor early during NAC with pCR in LABC patients. We also

studied both pre- and post-therapy association of FLT uptake with the tissue proliferative marker Ki-67 to compare ^{18}F-FLT-PET/CT against an accepted reference standard for cellular proliferation.

MATERIALS AND METHODS

Patients and Study Design

ACRIN 6688 was an observational, non-randomized, multicenter phase II study. This study was approved by the Institutional Review Board (IRB) of each participating center and a written and all subjects signed a written informed consent. The primary objective was to correlate the percent change in SUV_{max} between pre-therapy (FLT1) and ^{18}F-FLT-PET/CT after one cycle (FLT2) ($\%\Delta SUV_{max_FLT1\text{-}FLT2}$) of NAC with pCR in breast cancer patients for whom NAC was clinically indicated. The eligibility criteria included, histologically confirmed breast cancer diagnosis, primary breast cancer measuring greater than or equal to 2.0 cm, being a candidate for NAC and surgical resection of residual primary tumor after NAC, and no evidence of stage IV disease. The chemotherapy regimens chosen for each patient were not specified by trial design as long as the patients were on a treatment containing cytotoxic agent(s) as key components of the regimen. Hormonal and other targeted therapies were allowed, but only when given in association with chemotherapy agents. The eligible patients were planned to undergo three ^{18}F-FLT-PET/CT studies, FLT1, FLT2 and after completion of NAC before surgery (FLT3).

The key secondary objectives were to measure the correlation between FLT-1 and FLT-3 uptake parameters and immunohistochemical (Ki-67) analysis of pre-therapy biopsy and post-therapy surgical specimens, respectively. Other secondary objectives included in this report were to evaluate the relationship between FLT1, FLT2 and FLT3 uptake parameters and a) pCR of the primary tumor and residual cancer burden (RCB) (18), b) pCR to NAC in lymph node metastases; as well as to confirm safety and define adverse effects of FLT.

^{18}F-FLT **-PET/CT Protocol**

FLT was used under the authority of a National Cancer Institute (NCI) sponsored investigational new drug (IND) application. Following the injection of 2.6 MBq/kg (mean 167 MBq; range 110–204 MBq), a whole body image (5-7 bed positions) was obtained at 60 minutes (mean 70 min; range 50–101min). All patients were scanned on a calibrated and ACRIN-accredited PET/CT scanner, which included review of image quality and testing of SUVs using a uniform phantom and review of images as previously reported (19). A 60-minute dynamic regional PET/CT imaging was optional (results not included in this report), and if done the imaging sequence was to obtain the dynamic PET imaging first, then followed by the torso survey using static PET imaging. There were three planned ^{18}F-FLT-PET/CT sessions, baseline (FLT1) was completed within 4 weeks prior to NAC initiation. The early therapy (FLT2) scan was planned at 5-10 days after the initiation of the first cycle of the NAC and prior to the second cycle of NAC. Post-therapy FLT (FLT3) was performed after the completion of NAC and within 3 weeks prior to surgery. All sequential imaging sessions were performed on identical or technically equivalent PET/CT scanners for any individual patient. An adverse event evaluation was performed at each imaging time point. Following completion of NAC, the subjects underwent surgical resection of the breast primary (segmental or total mastectomy) and axillary nodal evaluation.

^{18}F-FLT **-PET/CT Image Data Analysis**

All ^{18}F-FLT -PET/CT images were transferred to the ACRIN Core Laboratory for quality control, archiving and analysis. Primary image interpretation was based on semiquantitative analysis (SUV) at a Core Laboratory site at the University of Washington. Image review and region placement was supervised by two nuclear medicine board-certified physicians with extensive experience in PET/CT, blinded to patient characteristics and outcome. Participating sites first indicated up to three primary tumor locations and up to five other non-primary tumors based upon local interpretation. At the Core Laboratory, volume of

interests (1cc ,VOIs) were positioned over the area of highest activity for both the primary and non-primary breast tumors at FLT1. The SUVpeak, defined as the average SUV value from a 1.0 cm diameter circular VOI (range 0.75-1.5 cm depending on scanner resolution) centered over the hottest tumor pixel at FLT1, was also created. VOIs were constructed on FLT2 and FLT3 images based upon CT localization, and residual tumor uptake when present, for all sites visualized at FLT1. The VOIs for FLT1, FLT2, FLT3 were verified independently by two expert reviewers blinded to the clinical and pathological results. The SUV_{peak}, as described above, the SUV_{mean}, the average SUV tumor value, and SUV_{max}, the maximal pixel intensity in the 1cc tumor VOI were recorded. For multiple primary tumors, the mean SUV_{max} for all tumor sites was used as the patient's overall tumor SUV. The axillary LNs were analyzed using the same methodology, including only those LNs measuring $\geq$1.0 cm in maximum dimension on CT to minimize partial volume effects. No partial volume corrections of the SUVs were attempted given the challenges of determining tumor boundaries from CT and the variability of the included scanner types.

Histopathology Analysis

The paraffin blocks or five unstained sections containing tumor tissue from the diagnostic biopsy and post-treatment surgical specimen were collected and sent to the Core Pathology Laboratory at Virginia Commonwealth University (CPL-VCU) for analysis. No additional biopsy was obtained. If there was residual tumor on post-treatment specimens, a representative section was acceptable but if no residual tumor was present per original pathology reports or in the post-treatment section, the entire tumor bed was sectioned to confirm pCR. If no viable or residual tumor remained on review of all sections, a pCR was documented.

Pathological Response. pCR is defined as absence of viable invasive tumor at histopathologic examination of post-therapy breast surgical specimen. Residual ductal carcinoma in situ (DCIS) in the absence of viable invasive cancer was considered a pCR; this was performed at the treating site and

reviewed at CPL-VCU. Dichotomous response assessment was performed with a result of either pCR or no-pCR.

Residual Cancer Burden Categories. A secondary measure assessed the RCB at CPL-VCU. The RCB was calculated as an index combining pathology measurements as described previously (18).
Four RCB categories: RCB-0 (pCR); RCB-1 (minimal residual disease); RCB-2 (moderate residual disease); RCB-3 (extensive residual disease or chemoresistant).
Lymph node status analysis: The participating sites' pathology reports were reviewed for LN status. In the case of a positive lymph node, a section of the LN was requested if size was unavailable from the report. In positive cases when a section was unavailable, if the original pathology report indicated macrometastasis, the size was assumed greater than 2 mm.

Ki-67 Analysis. An index of cellular proliferation was determined on pre- and post-treatment paraffin-embedded specimens (3-5 µm sections), by immunohistochemistry using DAKO monoclonal mouse antihuman Ki-67 antigen, (clone MIB-1) (DAKO Denmark A/S). A Ki-67 score was defined as the percentage of total number of tumor cells with nuclear staining over the total number of tumor cells in 10 high power fields (at 400X, Nikon Eclipse 80i light microscope)(20). Ki-67 labeling index (LI) was also calculated as the number of Ki-67 positive tumor cells per one thousand tumor cells.

Adverse Effect Assessment

An adverse event evaluation form was completed after all imaging time points. All adverse events were recorded within a 24-hour period.

Statistical Analysis

Data were analyzed using SAS v9.4 (SAS Institute, Cary, NC) and R v3.1.0 (R project, www.r-project.org). The study was designed to accrue 54 patients (including 10% dropout) to detect a difference of 0.25 in

AUC (area under the ROC curve) between the null hypothesis (AUC 0.50, i.e. by chance) and the alternative hypothesis (true AUC $\geq$0.75) with the significance level 0.05 and power 0.80. The pCR rate was assumed to be 0.25 in the sample size calculation. The percent change was defined as (SUV at FLT1 – SUV at FLT2)/SUV at FLT1 x 100 (% Δ $SUV_{max\_FLT1\text{-}FLT2}$). The pCR was based on the pathology results as described above. The empirical AUC was calculated and the 90% confidence interval was constructed from 2000 Bootstrapping. The optimal cutoff on the ROC curve was estimated using Youden's index method (21). The corresponding sensitivity and specificity were then calculated using the optimal cut-off for using the percent change to detect pathological complete response. Delong's method was used to test if the observed AUC was significantly greater than 0.5 with the one-sided p value (22). We also tested if SUV reduction in the pCR group was significantly larger than that in the no-pCR group using the Wilcoxon method with one-sided p value. A key secondary objective was to evaluate the correlation between SUV and Ki-67 LI at the baseline PET and at the PET after the treatment. This was quantified by the Spearman correlation coefficient (r) using the Cohen and Cohen method (23). For the LN evaluation, a three-category grouping was first implemented (i.e., 0 positive nodes, 1-3 positive nodes, and >3 positive nodes), and then Kruskal-Wallis one-way ANOVA was conducted to compare SUV distributions among groups with two-sided p values. For RCB evaluation, the four categories were binned into two groups (RCB 0/1 versus RCB 2/3) (18) and the comparison was conducted by Wilcoxon two-sample test with two-sided exact p value reported. The logistic regression was also fitted to quantify the association of the dichotomized RCB with % Δ $SUV_{max\_FLT1\text{-}FLT2}$ or % Δ $SUV_{max\_FLT1\text{-}FLT3}$.

RESULTS

A total of 90 patients were registered by 17 participating institutions (Appendix, Table I), between November, 2009 and August, 2012. All institutions had Institutional Review Board (IRB) approval of the protocol and all patients signed the informed consent form. Fifty-one of 90 patients were eligible for the primary objective analysis and completed both FLT1 and FLT2 scans within the study timeline. The

remaining 39 patients did not fulfill the primary aim eligibility criteria for various reasons (Appendix, Tables II and III). Up to 73 patients met the eligibility criteria for the secondary objective analyses correlating FLT uptake to the Ki-67, including 72 patients undergoing FLT1 and 43 undergoing FLT3. The patient characteristics are displayed in Table 1. All patients tolerated the ^{18}F-FLT-PET/CT protocol, and none suffered significant study related adverse effects.

Inherent to the protocol design, chemotherapy protocols and timing varied significantly among participants (Appendix, Table IV). There was some variability in uptake times, related to optional dynamic scans performed at several sites (not included in this analysis). However, uptake times from the same patient in serial scans were similar (Appendix, Table V). Treatment for over 60% of the patients including a combination of doxorubicin with cyclophosphamide, followed or preceded by a taxane.

Primary Objectives

^{18}F-FLT-PET/CT Imaging and pCR. Of 51 patients eligible for primary objective analysis, 43 completed all FLT1, FLT2, FLT3 imaging studies and 8 completed only FLT1 and FLT2 studies. All evaluable patients had measurable disease in the breast (>2.0 cm, median, 4.0 cm; range, 2.0-13.0). In large tumors, regions of imaging-based necrosis were excluded from the analysis. A pCR was achieved in 9 (18%) patients; of the remaining 42 (82%) patients who had a no-pCR, 31 (73%) had partial and 11 (26%) had no pathologic response. The median interval between FLT1 and initiation of chemotherapy was 3 days (range: 1-38), median interval between first chemotherapy and FLT2 was 7 days (range: 3-17) and in all cases FLT2 was prior to the second NAC cycle; median interval between FLT3 and surgery was 8 days (range: 1-70).

$\% \Delta SUV_{max\_FLT1\text{-}FLT\text{-}2}$ by pCR. Examples of serial FLT PET/CT studies are shown in Figures 1 and. The FLT1 SUV_{max} was not different between pCR and no-pCR groups [mean 6.1$\pm$3.2 (SD) vs. 5.6$\pm$3.0; difference -

0.5±3.0, 95%CI -2.7 - 1.7, p=0.62] (Table 2, Appendix, Table VI, VII). The $\%\Delta SUV_{max\_FLT1\text{-}FLT\text{-}2}$ changes are presented in Table 3. There was a marginal difference in $\%\Delta SUV_{max\_FLT1\text{-}FLT\text{-}2}$ between pCR and no-pCR patients (Wilcoxon, one-sided p=0.050). The corresponding value for $\%\Delta SUV_{peak\_FLT1\text{-}FLT\text{-}2}$ and $\%\Delta SUV_{mean\_FLT1\text{-}FLT\text{-}2}$ were similar (p=0.056, 0.033, respectively).

ROC Analysis of $\%\Delta SUV_{max\_FLT1\text{-}FLT2}$ in Predicting pCR. The AUC for $\%\Delta SUV_{max\_FLT1\text{-}FLT2}$ was 0.68 [(90%CI, 0.50,0.83), Delong's one-sided p= 0.046] (Figure 3), which is marginally significant from null hypothesis of AUC 0.50. The AUC for $\%\Delta SUV_{peak\_FLT1\text{-}FLT2}$ was 0.67 (90%CI, 0.50 – 0.82, Delong's one-sided p= 0.044). The AUC for $\%\Delta SUV_{mean\_FLT1\text{-}FLT2}$ was 0.70 (90%CI, 0.54 – 0.84, Delong's one-sided p= 0.016). The optimal cutoff of $\%\Delta SUV_{max\_FLT1\text{-}FLT2}$ was 51% from the Youden's index and the corresponding sensitivity and specificity were 0.56 (95%CI: 0.21, 0.86) and 0.79 (95%CI: 0.63, 0.90).

Secondary Objectives

$\%\Delta SUV_{max\_FLT1\text{-}FLT3}$ by pCR. Of 43 patients who had both FLT1 and FLT3 scans, the mean reduction in SUV_{max} ($\%\Delta SUV_{max\_FLT1\text{-}FLT3}$) was 67% [range: -4% - 96%] (Table 3). There was a significant difference in $\%\Delta SUV_{max\_FLT1\text{-}FLT3}$ between 8 pCR and 35 no-pCR patients (Wilcoxon, one-sided p= 0.0013). The corresponding values for $\%\Delta SUV_{mean\_FLT1\text{-}FLT3}$ and $\%\Delta SUV_{peak\_FLT1\text{-}FLT3}$ were also similar (not shown). The AUC for $\%\Delta SUV_{max\_FLT1\text{-}FLT3}$ in the prediction of pCR was 0.83 (90%CI, 0.72–0.94, Delong's one-sided p < 0.001] (Figure 4). AUC for $\%\Delta SUV_{peak\_FLT1\text{-}FLT3}$ was 0.82 (90%CI, 0.69–0.92, Delong's one-sided p< 0.001] and AUC for $\%\Delta SUV_{mean\_FLT1\text{-}FLT3}$ was 0.80 (90%CI, 0.68–0.92, Delong's one-sided p< 0.001].

Correlation Between FLT SUV_{max} and Ki-67 Expression. Suitable histopathology was available in 72 of the 90 recruited patients. There was a weak correlation between FLT1 SUV_{max} and Ki-67 (r =0.35, 95% CI 0.13, 0.54, p=0.002, Figure 5).

There were 43 patients who had suitable post-NAC tissue samples for correlation between surgical specimens and FLT3 SUVs. There was an improved correlation between FLT3 and Ki-67 (*r*= 0.68, 95% CI 0.47, 0.81, p<0.0001, Figure 6) compared to that of FLT1 and Ki-67 (p=0.020 for the difference in correlations). The correlation between $\%\Delta SUV_{max\_FLT1\text{-}FLT3}$ and $\%\Delta Ki\text{-}67_{FLT1\text{-}FLT3}$ was 0.57 (95% CI 0.35, 0.75 p < 0.0001) from the analysis of 42 patients who had Ki-67 expressions at both FLT1 and FLT3 time points.

Lymph Node Evaluation. Data on 38 patients were available for histopathological LN evaluation after NAC: 14 with negative nodes, 15 with 1-3 LN metastases and 9 with >3 LN metastases. The mean SUVmax was not different among three groups at all time-points FLT1, FLT2 and FLT3 (Appendix, Table VIII). In patients with negative nodes, mean FLT SUV_{max} at FLT1 and FLT3 were 4.4$\pm$3.0 (SD) and 0.8$\pm$0.4, respectively, compared to 6.6$\pm$3.6 and 1.2$\pm$0.7 for those with 1-3 positive LN after NAC (or to 7.3$\pm$5.0 and 2.5$\pm$2.2 for those with >3 positive nodes after NAC).

$\%\Delta SUV_{max\ FLT1\text{-}FLT\text{-}2}$ and LN Status at Surgery. The $\%\Delta SUV_{max\_FLT1\text{-}FLT2}$ and $\%\Delta SUV_{max\ FLT1\text{-}FLT3}$ were not associated with the LN status; the mean $\%\Delta SUV_{max\_FLT1\text{-}FLT2}$ and $\%\Delta SUV_{max\ FLT1\text{-}FLT3}$ were not different among negative (0 positive LNs), 1-3 positive and >3 positive groups (Appendix, Table IX).

^{18}F-FLT -*PET/CT and Primary Tumor RCB.* Suitable histopathology was available in 35 patients for RCB evaluation: 14 patients with RCB 0 or 1 and 21 patients with RCB 2 or 3. The SUV_{max} measurements at FLT1 and FLT2 were not different between these two groups (p=ns) while there was a difference in SUVmax at FLT3 (p=0.010) (Appendix, Table X). The $\%\Delta SUV_{max\_FLT1\text{-}FLT2}$ were not associated with RCB when RCB was dichotomized (p=ns) while $\%\Delta SUV_{max\_FLT1\text{-}FLT3}$ was (p<0.001); the mean values were different between the two groups at FLT3 (Appendix, TableXI). The magnitude of $\%\Delta SUV_{max\_FLT1\text{-}FLT3}$

tended to be associated with lower RCB when RCB was evaluated as a dichotomized variable (RCB 0-1 vs. RCB 2-3, (13) (odds ratio for $\%\Delta SUV_{max\_FLT1\text{-}FLT3}$: 0.86, 95%CI: 0.76-0.97, p=0.013).

Adverse Effects. There were no severe or life threatening events. In total, there were 20 adverse events; only 2 mild adverse events were considered to be possibly related to FLT injection; both were patient complaints of facial warmth during scanning.

DISCUSSION

Changes proliferative status of breast cancer measure by tissue assay early after chemotherapy or endocrine therapy are predictive of outcome (24). Furthermore, ^{18}F-FLT -PET, as a surrogate of cell proliferation, was shown to detect therapy induced proliferative changes as early as one week after chemotherapy (11, 13). In our study testing FLT PET as a predictor of tumor response to NAC after a single treatment dose, we found a marginal predictive value of $\%\Delta SUV_{max\_FLT1\text{-}FLT2}$ for pCR, with AUCs in the range of 0.66–0.70 for different measures of FLT uptake. The %ΔSUVs of FLT1 to FLT2 or FLT1 to FLT3 were not associated with the axillary LN status at surgery. We also found that post-therapy FLT uptake had a significant correlation with the Ki-67 proliferative index in post-therapy surgical specimens, further supporting FLT as a marker of tumor proliferation and therapeutic response. Additionally, the change between pre-therapy and post-therapy FLT PET (FLT1-FLT3) was a strong indicator of pCR, with AUCs of 0.80-0.83. However, while our results suggest an ability of early serial ^{18}F-FLT -PET/CT to predict therapeutic response, the accuracy of this predictive value was modest.

There has been variability in the published studies for ^{18}F-FLT -PET/CT as an early indicator of response. Small, single-center preliminary studies indicated good predictive value for response (11-14), while others did not (12, 15-17). Our multi-center study suggested some utility for FLT as an early predictor of

response, but preliminary reports of other multi-center trials did not find a predictive value for FLT (25). The difference between our study and the previous others likely lies with variations in end-points (clinical vs. pathologic response), image acquisition and analysis methods, patient populations, and therapy regimens. None of the prior studies investigated the potential of ^{18}F-FLT-PET imaging for predicting pCR to NAC, but rather investigated overall clinical response (11), changes in tumor size and in tumor markers in patients with all stages including distant metastasis (13). Our study had the advantage of having a slightly larger patient population, rigorous central system qualification, image and pathological analyses. While there was some variability in FLT uptake time in our study due to the optional dynamic imaging protocol, uptake times for each patient across longitudinal PET scans were reasonably consistent. The published results of multiple prospective studies indicated that early FDG-PET can reasonably predict pathological response to NAC in breast cancer (26-29). Our results on ^{18}F-FLT -PET were comparable with the prior published data by showing that NAC resulted in a significant SUV_{max} reduction (meanΔ, 39%;cut-off:51%) despite a marginal significance between pCR and non-pCR groups. However, the relatively small number of patients with a pCR as well as the varying definitions of the pCR across studies should be taken into consideration when comparing our results with others.

The main limitation of our study was that it did not control treatment types, and there was a considerable variation in chemotherapeutic regimens – both type and frequency of treatments – across participating centers, resulting in variable definitions of one treatment cycle. This likely contributes to variability in FLT results between patients and limits the overall performance of FLT PET as an early indicator of response in our study. Variability in scan timing and timing relative to treatment may also have contributed to variable performance. In our cohort, only 18% of patients had a pCR, a result that is consistent with other trials in LABC, using standard anthracycline–taxane combination chemotherapy in patients with heterogeneous tumor subtypes (30-32-). It is also important to realize the unbalanced number of patients between the pCR and no-pCR groups (9 vs. 42). The limited number of patients with

a pCR may have limited the ability to identify differences between high and low %ΔSUV between baseline and after one cycle of therapy. Further studies comparing changes in FLT to longer-term survival endpoint might yield further insights but was beyond the scope of this study. Similarly, a recent study by Woolf et al (15) reported that neither the baseline value nor the change in SUV_{max} after one cycle of NAC predicted treatment response, although most patients had a sizeable SUV_{max} reduction. Likewise, in our study, the mean decline in $SUV_{max\ FLT1\text{-}FLT2}$ was 39%. It should also be emphasized some data using tissue markers of proliferation suggest better predictive value in more aggressive tumor subtypes (e.g., HER-2 positive, basal or triple-negative cancers) and that these different tumor subtypes may have different responsiveness to cytoxic chemotherapy regimens (33, 34,). FDG-PET was shown to be most predictive in high-risk breast cancer phenotypes (4, 5). Our data were not conducive to perform subgroup analysis with respect to various risk groups, i.e. triple negative, HER-2 positive, because it was not powered for analyzing tumor subtypes. This analysis remains to be further pursued in future studies

As a secondary objective, correlation with Ki-67 revealed only a weak correlation between SUV_{max} at FLT1 and Ki-67 expression from pre-therapy biopsy specimens but a better correlation was found between FLT3 SUV_{max} and Ki-67 in post-therapy surgical specimens. Although prior studies found a correlation between FLT uptake and Ki67 expression (10, 13, 15), these results varied depending on both sample acquisition and post-acquisition treatment of tissues, possibly affecting direct comparisons. A recent breast cancer study demonstrated a significant difference between pre-therapy core biopsy and surgical sample Ki-67 values ($P< 0.0001$) in paired samples from untreated patients (35). Importantly, the difference represented an average difference in proliferation, with the core biopsies demonstrating a higher proliferation index compared to the surgical samples. Another factor potentially contributing to a variable pre-therapy relationship between FLT uptake and Ki-67 is the range of included tumor subtypes; some pre-clinical studies suggested that FLT uptake and its correlation with Ki-67 may vary

across tumor types (36). In our study, the weaker correlation in pre-therapy biopsy samples compared to surgical specimens may also be related to intra-tumor spatial heterogeneity of cell proliferation (37). A metaanalysis by Chalkidou et al showed that, the FLT uptake and Ki-67 correlation is significant and independent of cancer type (38). They also reported that the whole surgical specimen provided a significant correlation, while biopsy samples did not. This finding is in line with our results. The agreement between post-therapy FLT and Ki-67 supports the utility of FLT as a response marker, as does a more robust ability to discern pCR on the basis of differences between the pre- and post-therapy FLT uptake (AUC=0.80-0.83 for different SUV measures).

Some studies suggest that kinetic analysis of FLT uptake in breast cancer reportedly correlated better with Ki-67 than SUVs for distinguishing responders from non-responders early during chemotherapy (8, 39). We collected kinetic data in a select group of patients, but these results will be reported separately.

Residual disease is a continuous variable consisting of a range of tissue responses from complete response to refractory disease. The RCB index was proposed as a determinant of the extent of residual disease in the surgical specimens after NAC (18). The $\%\Delta SUV_{maxFLT1\text{-}FLT3}$ predicted RCB when RCB was evaluated as a continuous or dichotomized variable. The absolute SUV_{max} measurements at FLT-1 and FLT-2 were not different between the RCB groups; however, there was a significant difference in post-therapy FLT SUV_{max} (FLT3) for RCB0/1 vs. RCB2 or greater**.**

Conclusion. ^{18}F-FLT-PET/CT imaging of breast cancer after one cycle of NAC was marginally predictive of pCR despite highly variable chemotherapy regimens. Post-therapy NAC FLT uptake correlated with assay of Ki-67 on post-surgical tissue and with pCR. These early results, though not sufficient to support widespread use of FLT as an early response indicator for breast cancer, indicate potential efficacy of ^{18}F-

FLT T-PET/CT as an indicator of early therapeutic response of breast cancer to NAC and support future studies in a more uniformly treated patient population.

Acknowledgements: This work was supported by NCI Grants CA080098 (ARRA 2009), and U01 CA079778. The authors wish to acknowledge the efforts of the many centers that participated in this study, listed in Appendix, Table I. We also thank Glenna Gabrielli, data manager, ACRIN, Wei Bo, Biostatistics and Data Management Center Brown University, RI for their contributions.

REFERENCES

1. Chia S, Swain SM, Byrd DR, et al. Locally advanced and inflammatory breast cancer. *J Clin Oncol.* 2008;26:786-790.

2. Berruti A, Amoroso V, Gallo F, et al. Pathologic Complete Response as a potential surrogate for the clinical outcome in patients with breast cancer after neoadjuvant therapy: A meta-regression of 29 randomized prospective studies. *J Clin Oncol*. 2014;32:3883-3891.

3. Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet*. 2014;384:164-172.

4. Hatt M, Groheux D, Martineau A, et al. Comparison between 18F-FDG PET image-derived indices for early prediction of response to neoadjuvant chemotherapy in breast cancer. *J Nucl Med.* 2013;54:341-349.

5. Groheux D, Hindié E, Giacchetti S, et al. Early assessment with 18F-fluorodeoxyglucose positron emission tomography/computed tomography can help predict the outcome of neoadjuvant chemotherapy in triple negative breast cancer. *Eur J Cancer*. 2014;50:1864-1871.

6. Shields AF, Grierson JR, Dohmen BM, et al. Imaging proliferation in vivo with [F-18]FLT and positron emission tomography. *Nat Med.* 1998;4:1334–1336.

7. Bading JR, Shields AF. Imaging of cell proliferation: Status and prospects. *J Nucl Med*. 2008;49(suppl2):64S-80S.

8. Chang ZF, Huang DY, Hsue NC. Differential phosphorylation of human thymidine kinase in proliferating and M phase arrested human cells. *J Biol Chem*. 1994;269:21249–21254.

9. Munch-Petersen B, Cloos L, Jensen HK, Tyrsted G. Human thymidine kinase. I. Regulation in normal and malignant cells. *Adv Enzyme Regul*. 1995;35:69–89.

10. Kenny LM, Vigushin DM, Al-Nahhas A, et al. Quantification of cellular proliferation in tumor and normal tissues of patients with breast cancer by [18F] fluorothymidine-positron emission tomography imaging: evaluation of analytical methods. *Cancer Res.* 2005;65:10104–10112.

11. Kenny L, Coombes RC, Vigushin DM, et al. Imaging early changes in proliferation at 1 week post chemotherapy: A pilot study in breast cancer patients with3-deoxy-3-[18F]fluorothymidine positron emission tomography. *Eur J Nucl Med Mol Imaging.* 2007;34:1339-1347.

12. Contractor KB, Kenny LM, Stebbing J, et al. 18F]-3'Deoxy-3'-fluorothymidine positron emission tomography and breast cancer response to docetaxel. *Clin Cancer Res*. 2011;17:7664-7672.

13. Pio BS,Park CK, Pietras R, et al: Usefulness of 3'-[F-18]fluoro-3- deoxythymidine with positron emission tomography in predicting breast cancer response to therapy. *Mol Imaging Biol.* 2006;8:36-42

14. Marti-Climent JM, Dominguez-Prado I, Garcia-Velloso MJ, et al. [^{18}F]fluorothymidine-positron emission tomography in patients with locally advanced breast cancer under bevacizumab treatment: usefulness of different quantitative methods of tumor proliferation. *Rev Esp Med Nucl Imagen Mol.* 2014;33:280-285.

15. Woolf DK, Beresford M, Li SP, et al. Evaluation of FLT-PET-CT as an imaging biomarker of proliferation in primary breast cancer. *Br J Cancer.* 2014;110:2847-2854.

16. Contractor K, Aboagye EO, Jacob J, et al. Monitoring early response to taxane therapy in advanced breast cancer with circulating tumor cells and [(18)F] 3´-deoxy-3´-fluorothymidine PET: a pilot study. *Biomark Med.* 2012;6:231-233.

17. Couturier O, Rousseau C, Pierga J-Y, et al. 3'-deoxy-3'-[18F]fluoro-thymidine (18F-FLT) positron emission tomography (PET): An accurate and effective tool for assessing tumor response in breast cancer. Cancer Res 2013;73(24 Suppl):Abstract#P4-01-05.

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

19. Scheuermann JS, Saffer JR, Karp JS, Levering AM, Siegel BA. Qualification of PET scanners for use in multicenter cancer clinical trials: the American College of Radiology Imaging Network experience. *J Nucl Med.* 2009;50:1187-93.

20. Dowsett M, Nielsen TO, A'Hern R, et al. Assessment of Ki67 in breast cancer: Recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst*. 2011;103:1656-1664.

21. Youden, WJ. Index for rating diagnostic tests. *Cancer*. 1950;3:32-35.

22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-845.

23. Cohen J and Cohen P. *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.

24. Ellis MJ, Suman VJ, Hoog J, et al. Randomized phase II neoadjuvant comparison between letrozole, anastrozole, and exemestane for post- menopausal women with estrogen receptor-rich stage2 to 3 breast cancer:Clinical and biomarker outcomes and predictive value of the baseline PAM50-based intrinsic subtype–ACOSOG Z1031. *J Clin Oncol*. 2011;29:2342-2349.

25. Schmidt EV, Blackman S, Iannone R, et al. Limits of [18F]-FLT PET as a clinical biomarker of proliferation in breast cancer. SABCS 2013 Poster Session 4: Detection and Diagnosis: Molecular, Functional, and Novel Imaging (7:30AM-9:00AM):Friday, December 13, 2013,.

26. Rousseau C, Devillers A, Sagan C, et al. Monitoring of early response to neoadjuvant chemotherapy in stage II and III breast cancer by [18F]fluorodeoxyglucose positron emission tomography. *J Clin Oncol.* 2006;**24**:5366-5372.

27. Berriolo-Riedinger A, Touzery C, Riedinger JM, et al. [18F]FDG-PET predicts complete pathological response of breast cancer to neoadjuvant chemotherapy. *Eur J Nucl Med Mol Imaging.* 2007;**34**:1915-1924.

28. Schwarz-Dose J, Untch M, Tiling R, et al. Monitoring primary systemic therapy of large and locally advanced breast cancer by using sequential positron emission tomography imaging with [18F]fluorodeoxyglucose. *J Clin Oncol.* 2009;**27**:535-541.

29. Kolesnikov-Gauthier H, Vanlemmens L, Baranzelli MC, et al. Predictive value of neoadjuvant chemotherapy failure in breast cancer using FDG-PET after the first course. *Breast Cancer Res Treat* 2012;**131**:517-525.

30. Frasci G, D'Aiuto G, Comella P, et al. On behalf of the Southern Italy Cooperative Oncology Group (SICOG). Weekly cisplatin, epirubicin, and paclitaxel with granulocyte colony– stimulating factor support vs triweekly epirubicin and paclitaxel in locally advanced breast cancer: final analysis of a SICOG phase III study. *Br J Cancer*. 2006;95:1005–1012.

31. Villman K, Ohd JF, Lidbrink E, et al. A phase II study of epirubicin, cisplatin and capecitabine as neoadjuvant chemotherapy in locally advanced or inflammatory breast cancer. *Eur J Cancer*. 2007;43:1153–1160.

32. Ellis GK, Barlow WE, Gralow JR, et al. Phase III comparison of standard doxorubicin and cyclophosphamide versus weekly doxorubicin and daily oral cyclophosphamide plus granulocyte colony-stimulating factor as neoadjuvant therapy for inflammatory and locally advanced breast cancer: SWOG 0012. *J Clin Oncol.* 2011;29:1014-1021.

33. Carlomagno C, Perrone F, Gallo C, et al. c-erbB2overexpression decreases the benefit of adjuvant tamoxifen in early-stage breast cancer without axillary lymph node metastases. *J Clin Oncol.* 1996;14: 2702-2708.

34. Pritchard KI, Shepherd LE, O'Malley FP, et al. HER2 and responsiveness of breast cancer to adjuvant chemotherapy. *N Engl J Med*. 2006;354:2103-2111.

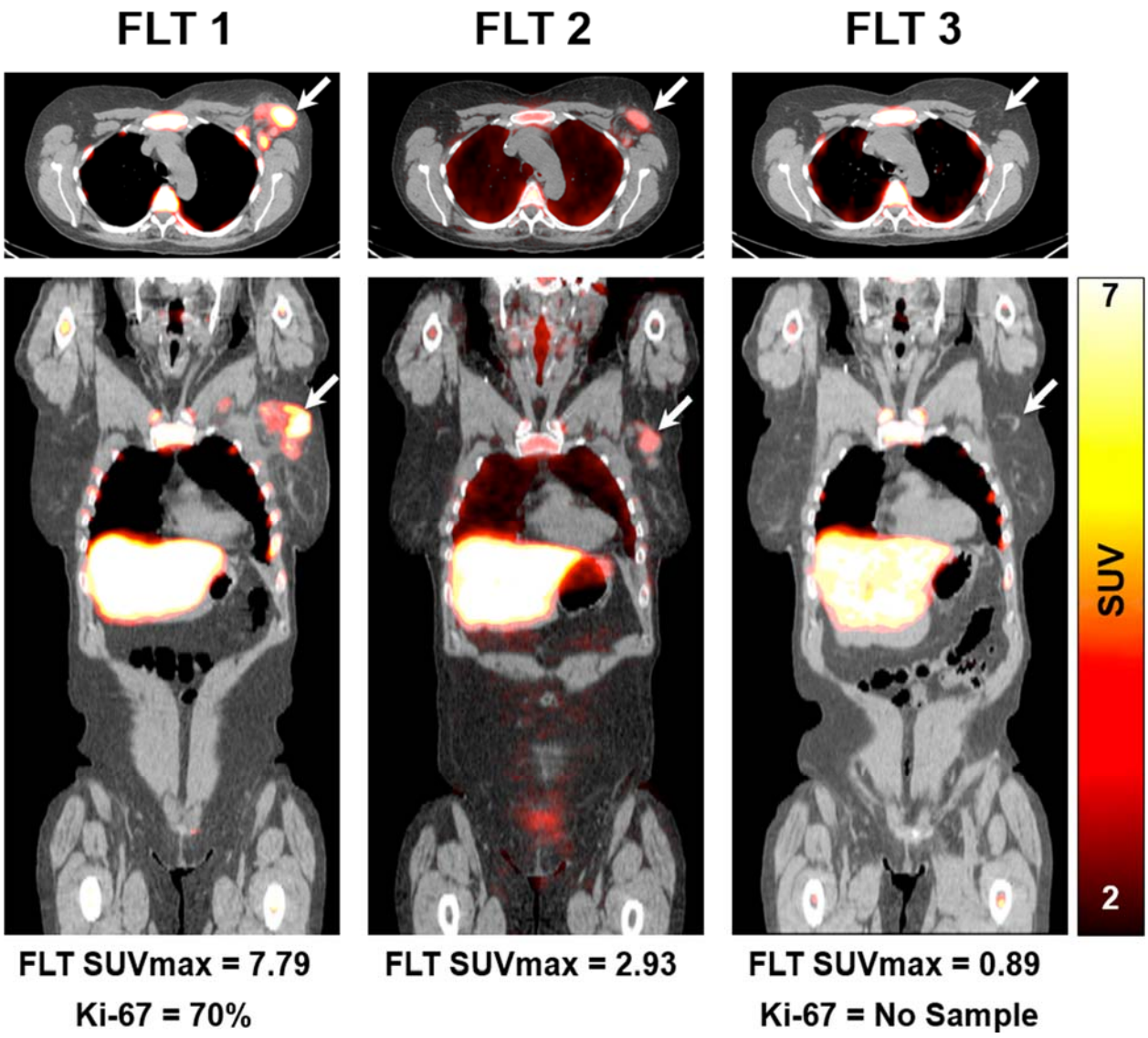
35. Romero Q, Bendahl PO, Klintman M, et al. Ki67 proliferation in core biopsies versus surgical samples—A model for neo-adjuvant breast cancer studies. *BMC Cancer*. 2011;11:341.

36. Zhang, Yan Z, Li W, Kuszpit K, et al. [(18)F]FLT-PET imaging does not always "light up" proliferating tumor cells. *Clin Cancer Res.* 2012 ;18:1303-1312.

37. Willaime JM, Turkheimer FE, Kenny LM, Aboagye EO. Quantification of intra-tumour cell proliferation heterogeneity using imaging descriptors of 18F fluorothymidine-positron emission tomography. *Phys. Med. Biol*. 2013;58:187-203.

38. Chalkidou A, Landau DB, Odell EW, et al. Correlation between Ki-67 immunohistochemistry and 18F-Fluorothymidine uptake in patients with cancer: a systematic review and meta-analysis. *Eur J Cancer*. 2012; 48: 3499–3513.

39. Gray KR, Contractor KB, Kenny LM, et al. Kinetic filtering of [(18)F] Fluorothymidine in positron emission tomography studies. *Phys Med Biol.* 2010;55:695-709 .



Figures 1. FLT PET/CT axial (upper panel) and coronal (lower panel) images demonstrate increased FLT uptake in an upper outer quarter breast tumor and an axillary lymph node, pre-therapy (left) with a substantial reduction in the primary breast tumor FLT uptake after 1 cycle of NAC (middle) and resolution of FLT uptake after completion of NAC (right). The patient had a pathologic complete response confirmed at surgery.

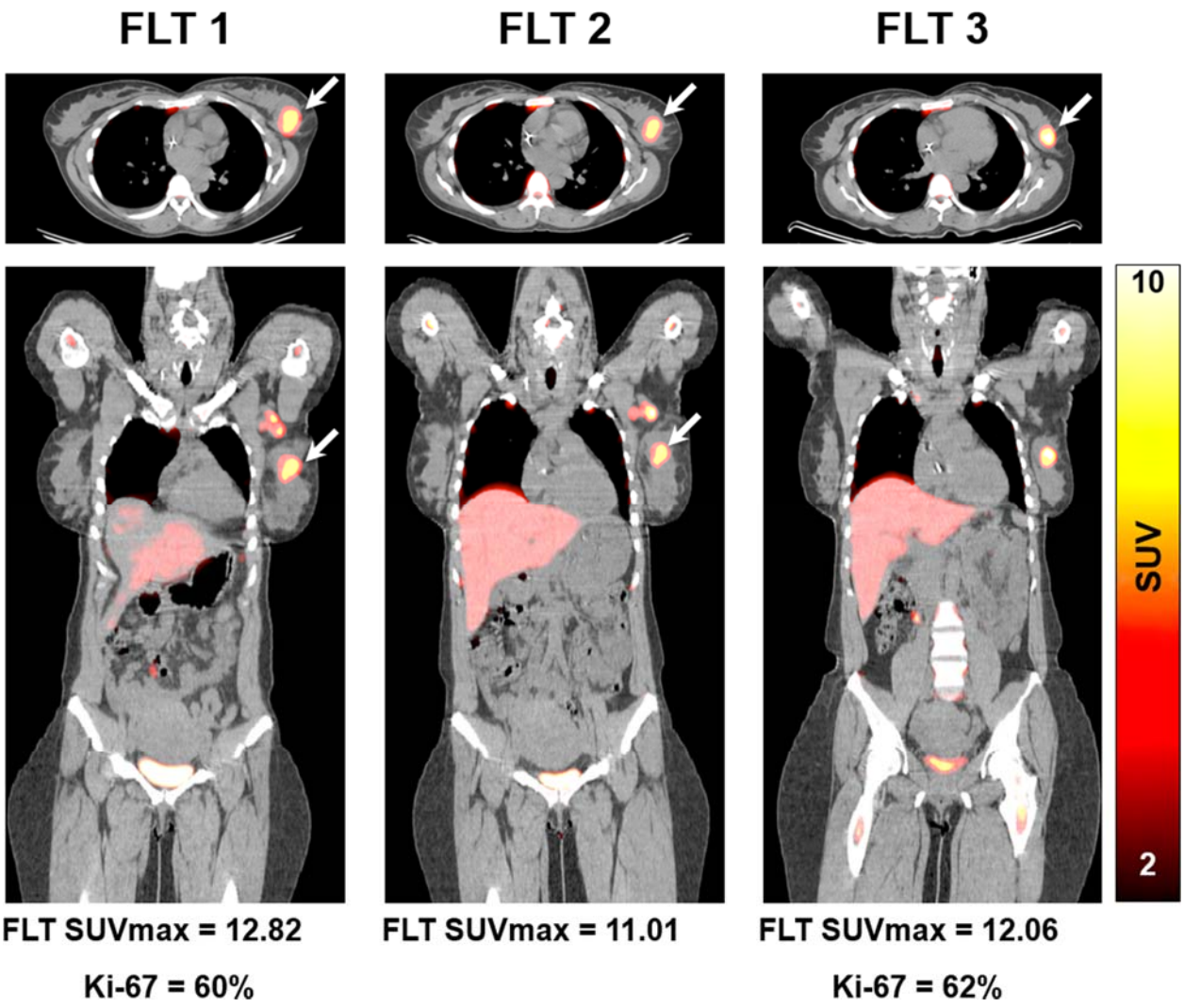


Figure 2. FLT PET/CT axial (upper panel) and coronal (lower panel) images demonstrate increased FLT uptake in an upper outer quarter breast tumor pre-therapy (left) with minimal decline in uptake and after one cycle of NAC (middle) and significant residual uptake after completion of NAC (right). At surgery significant residual viable tumor was confirmed (i.e., no-pCR) with a high Ki-67 index (62%).

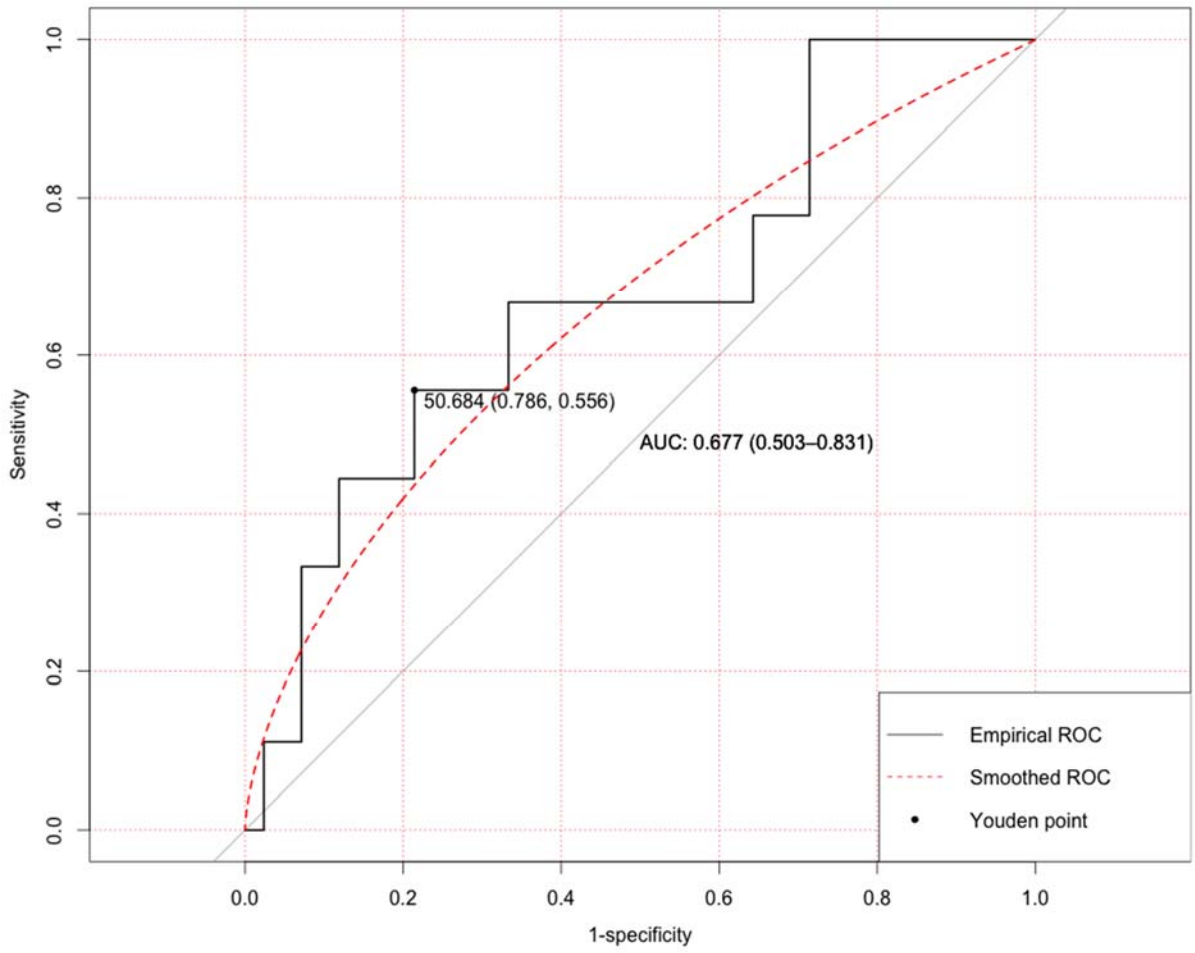


Figure 3. ROC curve of using $\%\Delta SUV_{max\ FLT1\text{-}FLT2}$ to predict pathologic complete response (pCR). The optimal cut point with corresponding specificity and sensitivity was identified through Youdan's index.

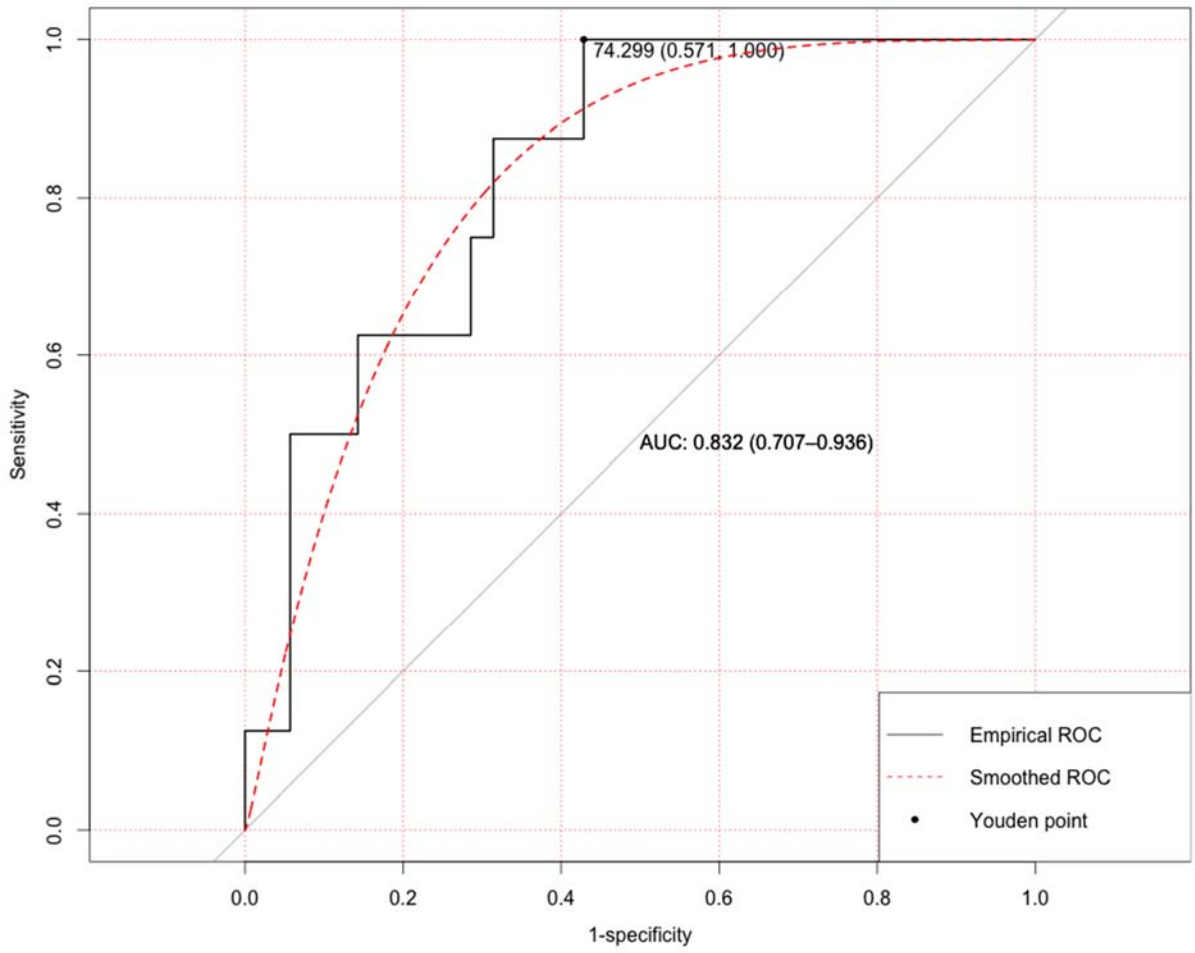


Figure 4. ROC curve of using $\%\Delta SUV_{max\ FLT1\text{-}FLT3}$ to predict pathologic complete response (pCR). The optimal cut point with corresponding specificity and sensitivity was identified through Youdan's index

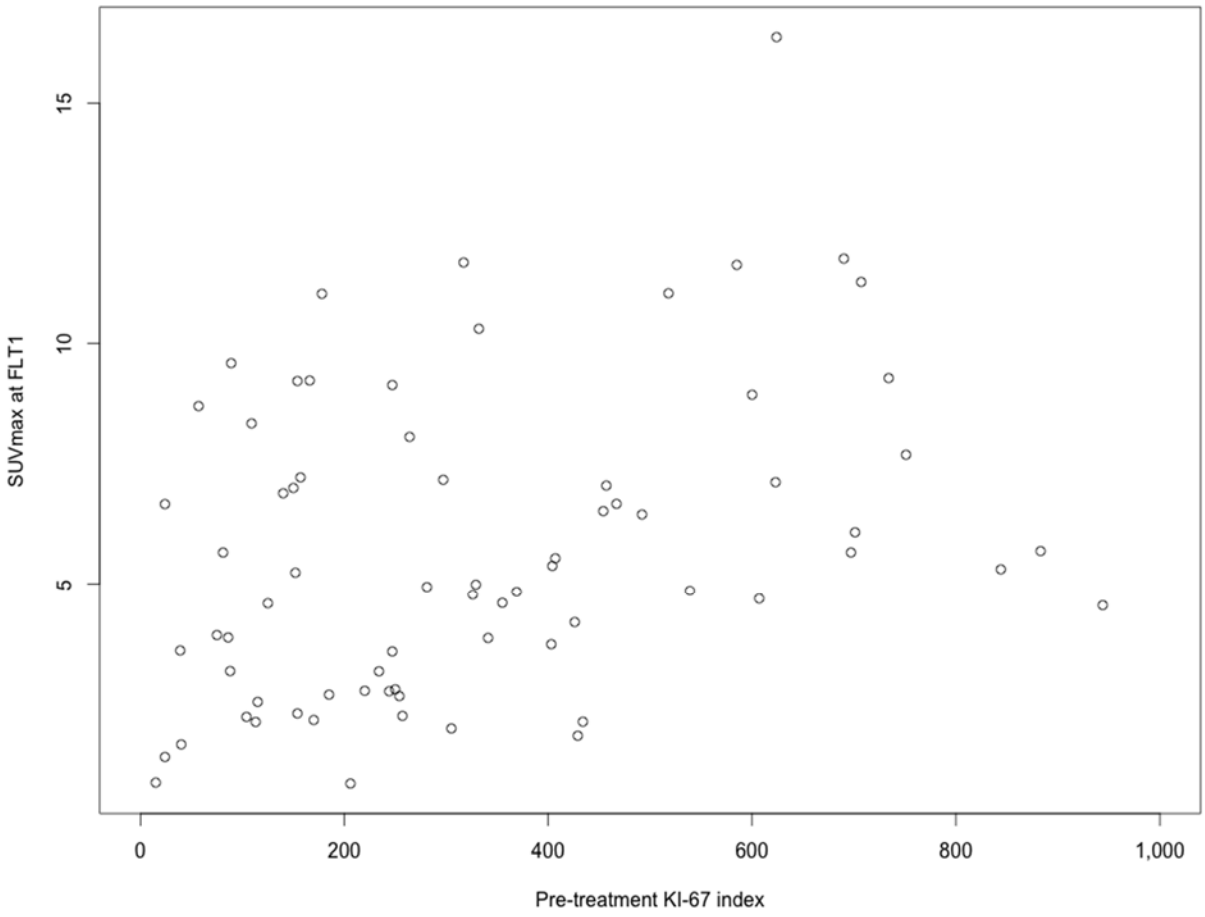


Figure 5. Scatter plot for Ki-67 on biopsy specimens versus SUVmax at FLT1.

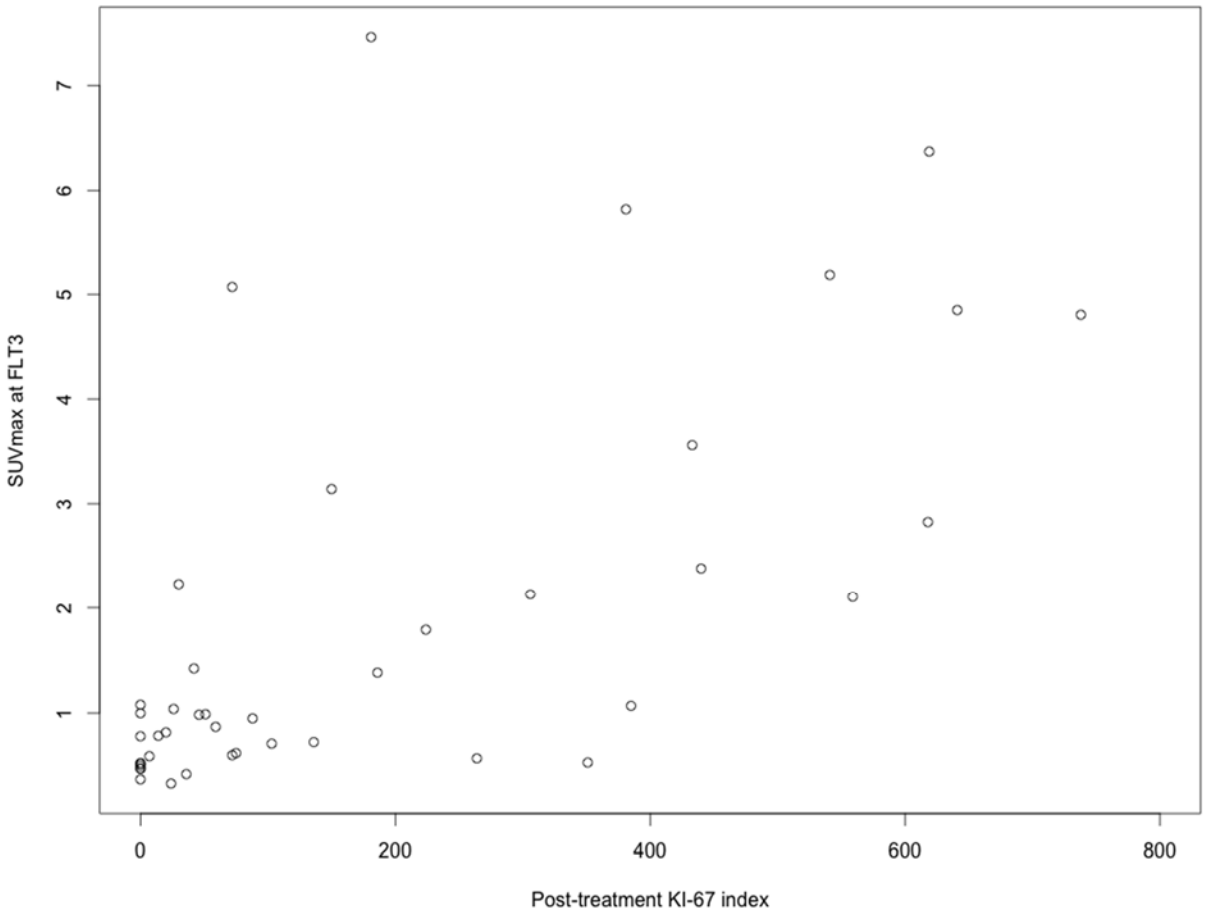


Figure 6. Scatter plot for Ki-67 on surgical specimens versus SUVmax at FLT3.

Table 1. Patient Demographics

Variable*	All Registered patients	Primary Aim analysis (n=51)	Ki-67 analysis (n=73)
AGE (yrs), mean (SD)	51.3 (10.9)	52.5 (10.6)	51.5 (10.2)
ETHNICITY			
HISPANIC OR LATINO	9 (10%)	3 (5.9%)	6 (8.2%)
NOT HISPANIC OR LATINO	75 (83.3%)	43 (84.3%)	62 (84.9%)
RACE			
AMERICAN INDIAN, ALASKA NATIVE, ASIAN	3 (3.3%)	2 (4%)	3 (4.1%)
BLACK OR AFRICAN AMERICAN	25 (27.8%)	12 (23.5%)	19 (26%)
WHITE	51 (56.7%)	32 (62.7%)	43 (58.9%)
TUMOR SIZE mean (CM), mean (SD)	4.4 (2.5)	4.6 (2.5)	4.4 (2.4)
MENOPAUSAL STATUS			
PRE-MENOPAUSAL	42 (46.7%)	22 (43.1%)	33 (45.2%)
POST-MENOPAUSAL	47 (52.2%)	29 (56.9%)	39 (53.4%)
INITIAL DIAGNOSIS			
INVASIVE BREAST CANCER NOS	2 (2.2%)	2 (3.9%)	2 (2.7%)
INVASIVE DUCTAL	77 (85.5%)	43 (84.4%)	63 (86.3%)
INVASIVE LOBULAR + MIXED INVASIVE&LOBULAR	10 (11.1%)	6 (11.8%)	8 (11%)
ER STATUS			
POSITIVE	49 (54.4%)	29 (56.9%)	43 (58.9%)
NEGATIVE	40 (44.4%)	22 (43.1%)	30 (41.1%)
PR STATUS			
POSITIVE	38 (42.2%)	20 (39.2%)	35 (47.9%)
NEGATIVE	51 (56.7%)	31 (60.8%)	38 (52.1%)
HER2 STATUS			
POSITIVE	32 (35.6%)	15 (29.4%)	22 (30.1%)
NEGATIVE	54 (60%)	34 (66.7%)	48 (65.8%)
RECEPTOR STATUS			
TRIPLE NEGATIVE	22 (24.4%)	13 (25.5%)	19 (26%)
OTHER	64 (71.1%)	36 (70.6%)	51 (69.9%)
T STAGE			
TX	3 (3.3%)	3 (5.9%)	3 (4.1%)
T1	1 (1.1%)	NA	NA
T2	42 (46.7%)	23 (45.1%)	35 (47.9%)
T3	31 (34.4%)	19 (37.3%)	25 (34.2%)
T4	12 (13.3%)	6 (11.7%)	10 (13.7%)
N STAGE			
pNX	3 (3.3%)	3 (5.9%)	3 (4.1%)
pN0	26 (28.9%)	13 (25.5%)	23 (31.5%)
pN1	45 (50.0%)	27 (53%)	35 (48.0%)
pN2	10 (11.1%)	5 (9.8%)	8 (11%)
pN3	5 (4.4%)	3 (5.9%)	4 (5.5%)
STAGE			
IIA	19 (21.1%)	10 (19.6%)	16 (21.9%)
IIB	32 (35.6%)	18 (35.3%)	26 (35.6%)
IIIA	22 (24.4%)	14 (27.5%)	18 (24.7%)
IIIB	9 (10%)	5 (9.8%)	8 (11%)
IIIC	4 (4.4%)	3 (5.9%)	4 (5.5%)
IV	2 (2.2%)	NA	NA
GRADE AT DIAGNOSIS			
1	2 (2.2%)	1 (2%)	2 (2.7%)
2	19 (21.1%)	12 (23.5%)	17 (23.3%)
3	44 (48.9%)	27 (52.9%)	33 (45.2%)

*The percentages not adding up to 100% are due to missing data; Tumor size was determined using baseline imaging;
ER, estrogen receptor; PR, progesterone receptor; pN, pathologic N stage

Table 2. Distributions of SUVmax and Ki-67 across different time points.

Time point	Aim	Parameter Tested	pCR Status	Number of Evaluable Participants	Range	Mean	SD	P value*
Baseline (FLT1)	Primary Aim	SUVmax	all data	51	0.87-11.76	5.65	2.97	
			pCR	9	1.86-11.76	6.09	3.17	0.62
			no-pCR	42	0.87-11.68	5.55	2.95	
After one cycle of NAC (FLT2)	Primary Aim	SUVmax	all data	51	0.44-8.29	3.21	1.84	
			pCR	9	0.92-8.29	3.00	2.54	0.35
			no-pCR	42	0.44-6.43	3.25	1.69	
After completion of NAC (FLT3)	FLT3 Secondary Aim	SUVmax	all data	43	0.33-7.46	1.74	1.81	
			pCR	8	0.37-1.08	0.62	0.27	0.0047
			no-pCR	35	0.33-7.46	2.00	1.92	
	Ki-67 Secondary Aim	Ki-67	all data	43	0-738	184.14	220.16	
			pCR	8	0-0	0.00	0.00	<0.001
			no-pCR	35	0-738	226.23	223.77	

SUV, standardized uptake value; pCR, pathologic complete response; *Two-sided exact p value from Wilcoxon two-sample test.

Table 3. Distributions of SUV differences between time points and by pathologic response status

Difference	pCR Status	Number of Evaluable Patients	Range	Mean (SD)	P value*
$\%SUV_{max\ FLT1\text{-}FLT2}$	all data	51	-30.2-86.7	38.8 (26.1)	
	pCR	9	24.7-84.3	52.7 (22.8)	0.050
	no-pCR	42	-30.2-86.7	35.8 (26.0)	
$\%SUV_{max\ FLT1\text{-}FLT3}$	all data	43	-4.04-96.1	66.9 (25.9)	
	pCR	8	74.5-96.1	86.9 (7.5)	0.0013
	no-pCR	35	-4.0-95.1	62.3 (26.5)	

SUV, standardized uptake value; pCR, pathologic complete response; *one-sided exact p value from Wilcoxon two-sample test.