18F-FDG PET/CT for Monitoring of Treatment Response in Breast Cancer

Stefanie Avril1, Raymond F. Muzic, Jr.2, Donna Plecha2, Bryan J. Traughber3, Shaveta Vinayak4, and Norbert Avril2

1Department of Pathology, University Hospitals Case Medical Center, Case Western Reserve University, Cleveland, Ohio; 2Department of Radiology, University Hospitals Case Medical Center, Case Western Reserve University, Cleveland, Ohio; 3Department of Radiation Oncology, Seidman Cancer Center, University Hospitals Case Medical Center, Case Western Reserve University, Cleveland, Ohio; and 4Department of Medicine (Hematology and Oncology), Seidman Cancer Center, University Hospitals Case Medical Center, Case Western Reserve University, Cleveland, Ohio

Changes in tumor metabolic activity have been shown to be an early indicator of treatment effectiveness for breast cancer, mainly in the neoadjuvant setting. The histopathologic response at the completion of chemotherapy has been used as the reference standard for assessment of the accuracy of 18F-FDG PET in predicting a response during systemic treatment. Although a pathologic complete response (pCR) remains an important positive prognostic factor for an individual patient, a recent metaanalysis could validate pCR as a surrogate marker for patient outcomes only in aggressive breast cancer subtypes. For establishment of the clinical application of metabolic treatment response studies, larger series of specific breast cancer subtypes—including hormone receptor–positive, human epidermal growth factor receptor 2–positive, and triple-negative breast cancers—are necessary. In addition, thresholds for relative changes in 18F-FDG uptake to distinguish between responding and nonresponding tumors need to be validated for different systemic treatment approaches, with progression-free survival and overall survival as references. A PET-based treatment stratification is applicable clinically only if valid alternative therapies are available. Of note, patients who do not achieve a pCR might still benefit from neoadjuvant therapy enabling breast-conserving surgery. In the metastatic setting, residual tumor metabolic activity after the initiation of systemic therapy is an indicator of active disease, whereas a complete resolution of metabolic activity is predictive of a successful treatment response.

Key Words: 18F-FDG; 18F-FDG PET; PET/CT; breast cancer; treatment response; therapy monitoring

DOI: 10.2967/jnumed.115.157875

Modern treatment strategies are tailored to molecular subtypes of breast cancer, allowing for a more individualized treatment approach. A pathologic complete response (pCR) is an important prognostic parameter and has commonly been used as a surrogate marker for a treatment response. Achievement of a pCR has been correlated with an improved long-term outcome, although only for aggressive breast cancer subtypes. Changes in tumor size represent an accepted endpoint for evaluating therapeutic effects in metastatic breast cancer. RECIST was established 15 y ago and has been updated since then. However, several cycles of treatment are often required before CT or MRI can detect a measurable change in tumor size.

A decrease in tumor metabolic activity offers both assessment of a treatment response after the completion of therapy and early prediction of therapeutic effectiveness after the first or second cycle of chemotherapy. Identifying nonresponding patients on the basis of changes in tumor metabolic activity early during treatment could facilitate a change from an ineffective to a more effective treatment approach. To allow the use of 18F-FDG PET–based treatment stratification in clinical practice, several items need to be addressed; these include the best timing for measuring changes in tumor metabolic activity during treatment and defined cutoff values for changes in tumor metabolic activity.

18F-FDG PET FOR EARLY PREDICTION OF TREATMENT RESPONSE

The first observation of early changes in tumor glucose metabolism of breast cancer occurred in 1993, when Wahl et al. described an early decrease in the metabolic activity of tumors responding to a combination of chemotherapy and endocrine therapy. Similar results were found in further studies often including small sample sizes. Smith et al. reported a significantly greater reduction in 18F-FDG uptake in patients who subsequently achieved a macroscopic pathologic response. Rousseau et al. studied 64 stage II and III breast cancer patients at multiple cycles during neoadjuvant chemotherapy and found a marked decrease in 18F-FDG uptake in nearly all patients who achieved a greater than 50% therapeutic effect. Schwarz-Dose et al. confirmed, in 104 patients, that the greater the reduction in tumor metabolic activity early during neoadjuvant treatment, the more likely the patients would achieve a pathologic response. After the first cycle of chemotherapy, tumor metabolic activity decreased by 50% ± 18% in pathologic responders; in comparison, the decrease in pathologic nonresponders was 36% ± 20%. Of note, all breast carcinomas with a baseline SUV of less than 3.0 did not respond to chemotherapy. In another
study, in 126 patients, a significantly higher baseline SUV (mean, 10.5) was found in 41 patients who subsequently achieved a pathologic response (defined as a reduction in tumor cellularity of >90%); in comparison, the baseline SUVs in partial responders and nonresponders were 6.9 and 5.2, respectively (12). A cutoff of an SUV of greater than 5.9 at baseline predicted a pathologic response with 78% sensitivity and 65% specificity.

Performing 18F-FDG PET after the second cycle of treatment potentially provides a more accurate prediction of a treatment response. Using a 40% decrease in the SUV, Rousseau et al. identified a negative predictive value of 68% for identifying nonresponders after the first cycle; this value increased to 85% after the second cycle (10). Schwarz-Dose et al. found, for histopathologic nonresponders, negative predictive values of 89.5% after the first cycle (cutoff: 45% decrease in the SUV) and 88.9% after the second cycle (cutoff: 55%); these findings indicated similar accuracies for predicting a nonresponse after the first and second cycles (11). A recent metaanalysis including 19 studies with more than 900 patients found that the best cutoff for a response was a decrease in 18F-FDG uptake ranging from 55% to 65% (13). Although the sensitivity and the specificity for identifying patients responding to treatment were limited (84% and 66%, respectively), the negative predictive value for identifying nonresponders was high (91%). Figure 1 shows a good treatment response after 2 cycles of chemotherapy in a right breast mass, whereas Figure 2 shows no metabolic response. A summary of these studies and the accuracy of the cutoff values are shown in Table 1.

The fact that the histopathologic reference standards used in these studies were based on different criteria limits a direct comparison. Most used a pCR as the reference standard; 8 of 19 studies also included patients with minimal residual disease as histopathologic responders (13). To make matters more complex, previous clinical trials were heterogeneous with regard to systemic treatment approaches, which included combinations of various chemotherapeutic agents and possible additional targeted treatments. Most neoadjuvant chemotherapy regimens were based on a combination of anthracyclines and concurrent or sequential administration of taxanes, often with the addition of cyclophosphamide or fluorouracil as a third agent (14). There is little information about the relationship between potential changes in tumor metabolic activity and specific treatment regimens, a factor that should be considered if 18F-FDG PET is to be used for treatment stratification. Another important consideration is tumor shrinkage in patients without a pCR, which would allow for breast-conserving surgery. This endpoint was not addressed in previous studies.

Few studies have assessed early changes in metabolic activity in axillary lymph nodes (9,15,16). A marked decrease in the SUV was observed after the first cycle of treatment in 52 patients achieving a nodal pCR (17). When a cutoff of a 50% decrease was used, 18F-FDG PET predicted a lymph node response with a sensitivity of 96%, a specificity of 75%, and a negative predictive value of 95%. Not all patients had cytologically confirmed lymph node metastases at baseline, a fact that limited the calculation of true responders. Of note, there was no correlation between a histopathologic response of the primary tumor and axillary lymph node metastases (17).

Hybrid PET/MRI technology was recently introduced, appearing in the clinical setting in 2007 (18,19). There is little literature regarding the use of 18F-FDG PET/MRI in breast cancer (20,21), as no prospective study has addressed the potential role for treatment monitoring. It would be of interest to directly compare changes in tumor metabolic activity with dynamic contrast enhancement or diffusion-weighted imaging. Future applications might include a combination of such parameters, as shown for other tumor entities (22).

HISTOPATHOLOGIC RESPONSE

There is strong evidence that a histopathologic assessment after the completion of neoadjuvant therapy is a surrogate marker for a
treatment response. Multiple large neoadjuvant trials have demonstrated a significant correlation between pCR and improved patient outcomes (2,23–25). Various definitions of a histopathologic response, based on the extent of residual invasive carcinoma in the breast and regional lymph nodes, regardless of residual in situ carcinoma, have been used (26). Of note, these various definitions of a histopathologic response have been shown to be significantly associated with patient outcomes. Therefore, PET trials with different histopathologic response criteria remain valid, although a comparison with SUV cutoffs is limited.

The largest metaanalysis to date was published in 2014 and included almost 12,000 women (27). Despite the strong prognostic information provided by a pCR in an individual patient, a pCR is not sufficient to demonstrate the superiority of a given treatment regimen over another, and survival data are still necessary for such comparisons. A proposal by an international working group recommended that a pCR should be defined as the absence of residual invasive carcinoma, with or without residual in situ carcinoma in the breast [ypT0/is ypNX or ypT0/is ypN0] (26,28). The panel also recognized a need for classification of the amount of residual tumor burden, which allows the measurement of residual tumor as a continuous variable, in addition to the dichotomized distinction of a pCR versus no pCR. The MD Anderson Cancer Center group has published criteria for measuring the extent of residual breast cancer burden, which has been shown to be associated with survival (26,28).

**PET THERAPY MONITORING FOR DIFFERENT BREAST CANCER SUBTYPES**

Molecular studies have identified intrinsic breast cancer subtypes characterized by common gene expression profiles (29,30). Immunohistochemical analysis of the estrogen receptor, the progesterone receptor, and human epidermal growth factor receptor 2 (HER2) is often used instead of gene expression to define 3 major subtypes of breast cancer: luminal (estrogen receptor–positive, progesterone

### TABLE 1

Summary of Studies of 18F-FDG PET for Early Prediction of Treatment Response

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>No. of patients included in 18F-FDG PET analysis</th>
<th>Pathologic response criteria*</th>
<th>Percentage of responders</th>
<th>18F-FDG PET metabolic response</th>
<th>Decrease in SUV</th>
<th>Percentage sensitivity</th>
<th>Percentage specificity</th>
<th>Percentage negative predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schelling et al. (8)</td>
<td>2000</td>
<td>22</td>
<td>pCR + pMRD</td>
<td>29</td>
<td>First, second</td>
<td>&gt;45% after first or second</td>
<td>100 (first), 83 (second)</td>
<td>85 (first), 94 (second)</td>
<td>N/A</td>
</tr>
<tr>
<td>Smith et al. (9)</td>
<td>2000</td>
<td>30</td>
<td>pCR-macro</td>
<td>38</td>
<td>First</td>
<td>&gt;20%</td>
<td>90</td>
<td>74</td>
<td>N/A</td>
</tr>
<tr>
<td>Rousseau et al. (10)</td>
<td>2006</td>
<td>64</td>
<td>&gt;50% therapeutic effect</td>
<td>56</td>
<td>First, second</td>
<td>&gt;40% after first or second</td>
<td>61 (first), 89 (second)</td>
<td>96 (first), 95 (second)</td>
<td>68 (first), 85 (second)</td>
</tr>
<tr>
<td>Schwarz-Dose (11)</td>
<td>2009</td>
<td>104</td>
<td>pCR + pMRD</td>
<td>16</td>
<td>First, second</td>
<td>&gt;45% after first, &gt;55% after second</td>
<td>73</td>
<td>63</td>
<td>90</td>
</tr>
<tr>
<td>Humbert et al.† (32)</td>
<td>2012</td>
<td>37</td>
<td>pCR²</td>
<td>38</td>
<td>First</td>
<td>&gt;75%</td>
<td>64</td>
<td>83</td>
<td>79</td>
</tr>
<tr>
<td>Groheux et al.† (35)</td>
<td>2013</td>
<td>30</td>
<td>pCR²</td>
<td>53</td>
<td>Second</td>
<td>&gt;62%</td>
<td>86</td>
<td>63</td>
<td>N/A</td>
</tr>
<tr>
<td>Humbert et al.† (34)</td>
<td>2014</td>
<td>54</td>
<td>pCR²</td>
<td>41</td>
<td>First</td>
<td>&gt;60%</td>
<td>83</td>
<td>52</td>
<td>84</td>
</tr>
</tbody>
</table>

SUV = 3.0³

SUV < 2.1³

* pCR = pathologic complete response, defined as absence of residual invasive tumor in breast, irrespective of lymph node status [ypT0/is ypNX], unless otherwise indicated; pMRD = minimal residual disease, defined as a few scattered foci of microscopic residual invasive tumor (<2 mm); pCR-macro = pathologic response, defined as absence of macroscopically visible tumor.

† pCR was defined as absence of residual invasive tumor in both breast and axillary lymph nodes [ypT0/is ypN0].

‡ Results were based on HER2-positive breast cancers only.

§ Absolute SUVs for defining cutoff values.

N/A = not available.
receptor–positive, or both, HER2-negative), HER2-positive, and triple-negative (estrogen receptor–negative/progesterone receptor–negative/HER2-negative) (31). HER2-positive and triple-negative tumors are generally more aggressive than luminal breast cancers.

An analysis of 115 women identified the highest SUVs (11.3 ± 8.5) in triple-negative tumors (32). The decrease in the SUV after the first cycle of systemic therapy was significantly higher in triple-negative and HER2-positive subtypes than in the luminal subtype. However, the decrease in the SUV was a predictor for a subsequent pCR only in HER2-positive tumors (accuracy, 76%). The molecular heterogeneity (33) of triple-negative tumors and their small number may partly explain the lack of a significant correlation with a pCR. Triple-negative breast cancers tend to have an aggressive clinical course but often respond to anthracycline- or taxane-based chemotherapy. Patients not achieving a pCR after neo-adjuvant treatment have a higher risk for early recurrences and shorter survival. The addition of platinum-based chemotherapy is a potential option, but it involves significant additional toxicity, and it is unclear which patients would benefit the most. These circumstances present an exciting opportunity to assess the role of 18F-FDG PET imaging in the prediction of a response.

In 57 HER2-positive patients treated with chemotherapy and trastuzumab, an SUV of less than 2.1 after the first cycle was the best independent predictor of a pCR (34). A decrease in the SUV of greater than 60% had the highest negative predictive value for identifying nonresponding HER2-positive breast cancers. Groheux et al. found, in 30 HER2-positive patients, that low residual 18F-FDG uptake (SUV, <3.0) after the second cycle was the best predictor of a pCR (35). A decrease in 18F-FDG uptake of 62% or more was also predictive of a pCR, but at a lower overall accuracy than absolute SUVs (73% and 90%, respectively). There are 2 competing analysis approaches—measuring absolute SUVs and measuring relative changes in SUVs—and no conclusion regarding which approach offers better identification of nonresponding tumors can be made on the basis of current literature.

Luminal (hormone receptor–positive, HER2-negative) breast cancers are characterized by lower metabolic activity, and patients with these cancers rarely achieve a pCR. A recent study with patient outcomes as a reference found that a poor metabolic response (<16% decrease in the SUV) after the first cycle was associated with a shorter 5-y overall survival relative to the findings for metabolic responders (49% vs. 96%) (36). Only 42 of 61 tumors (69%) were hypermetabolic at baseline and could be assessed with 18F-FDG PET; this factor represents a distinct limitation for this tumor subtype. Of note, patients with low tumor glucose metabolism at baseline had the best 5-y survival (100%). Groheux et al. studied 82 patients with hormone receptor–positive (HER2-negative) breast cancer and found that a small decrease in the SUV (<12% after the second cycle) was significantly associated with short event-free survival (37).

18F-FDG PET THERAPY MONITORING FOR METASTATIC BREAST CANCER

For patients with metastatic disease, there are several chemotherapy agents (including anthracyclines, taxanes, gemcitabine, and capecitabine) as well as targeted endocrine or anti-HER2 therapy. Treatment response is based on changes in tumor size, as a histopathologic assessment is often not practical or even possible. In accordance with RECIST (5), up to 5 measurable target lesions representative of involved organs are evaluated (5). An important limitation is that changes in tumor size often do not correlate with patient outcomes.

A pilot study enrolling 11 patients with 26 metastatic lesions revealed a statistically significant reduction in tumor metabolic activity after the first and second cycles of first-line chemotherapy in lesions that responded (38). The overall survival of nonresponding patients was significantly shorter than that of responding patients (8.8 vs. 19.2 mo). Patients not responding to treatment were identified several weeks earlier with 18F-FDG PET than with conventional imaging.

Eighty-two HER2-positive patients underwent dual targeted anti-HER2 therapy with lapatinib and trastuzumab (39). A metabolic nonresponse, defined as a decrease in the SUV of less than 25% after 1 wk, had a high negative predictive value (91%) for identifying patients who would not achieve an objective response according to RECIST. In addition, patients identified as metabolic nonresponders after week 1 had a shorter time to progression than responding patients. In 20 patients with metastatic breast cancer, a decrease in the SUV of greater than 45% after the third cycle was significantly associated with a clinical response at the completion of chemotherapy and a longer overall survival (40).

The use of early changes in tumor metabolic activity is more difficult for hormone receptor–positive breast cancer patients receiving antihormonal therapies. Studies revealed an increase in 18F-FDG uptake 7–10 d after the initiation of endocrine therapy. This “metabolic flare phenomenon” occurring within the first 1 or 2 wk of endocrine therapy was attributed to the initial agonist effect of tamoxifen and was found to be predictive of a positive response to therapy (41,42). Recently, 18F-FDG PET predicted progression-free survival in 22 patients with metastatic hormone receptor–positive breast cancer (43).

18F-FDG PET FOR RESPONSE ASSESSMENT AFTER COMPLETION OF THERAPY

Residual tumor metabolic activity after the completion of therapy is an indicator of residual viable tumor tissue, whereas a complete resolution of increased metabolic activity provides a high positive predictive value for a successful treatment response. In a neoadjuvant multicenter trial, 99 patients underwent 18F-FDG PET, mammography, ultrasound, and MRI before surgery (44). Patients who achieved a pCR had significantly lower 18F-FDG uptake than nonresponding patients. Nevertheless, the sensitivity of 18F-FDG PET for detecting residual tumor was only 32.9% when an SUV threshold of 2.0 was used; the sensitivity increased to 57.5% when a threshold of 1.5 was used. Conventional imaging modalities were more sensitive than 18F-FDG PET for identifying residual tumor but had lower specificity. Neither 18F-FDG PET nor conventional imaging could exclude the presence of residual viable tumor; this factor is an important limitation for all imaging modalities.

18F-FDG PET was performed after 3 cycles of high-dose chemotherapy in 47 metastatic breast cancer patients, and a negative post-treatment 18F-FDG PET result was the most powerful predictor of survival—superior to CT imaging (45). A total of 34 patients (72%) achieved a complete metabolic response and had a median survival of 24 mo; in comparison, the median survival of patients with a positive 18F-FDG PET result was 10 mo. According to multivariate analysis, the relative risk of death was highest in patients with 18F-FDG PET-positive disease (relative risk, 5.3).

Changes in the sizes of bone metastases are particularly difficult to evaluate with conventional imaging, as sclerotic lesions do not
disappear and lytic lesions can show sclerotic changes as an indicator of a treatment response. Two studies demonstrated a high sensitivity of $^{18}$F-FDG PET/CT for the detection of osseous metastases in patients with newly diagnosed metastatic breast cancer, and the metabolic activity of osseous breast cancer metastases provided prognostic information (46,47). Stafford et al. studied 24 patients with bone-dominant metastatic breast cancer and reported a correlation between changes in $^{18}$F-FDG uptake and the overall clinical assessment of a response (48). The same group’s subsequent study of 28 metastatic breast cancer patients revealed that patients with no change in $^{18}$F-FDG uptake were twice as likely to progress as those who showed a metabolic response (49). In a retrospective analysis, bone metastases in 102 patients were assessed with $^{18}$F-FDG PET/CT before and after treatment, and a decrease in $^{18}$F-FDG uptake was a significant predictor of the response duration in univariate and multivariate analyses (50). Treatment monitoring of bone metastases is discussed further in this supplement.

CONCLUSION

On the basis of data published to date, the monitoring of systemic treatment by $^{18}$F-FDG PET is a powerful tool for the early prediction and assessment of a response in patients with newly diagnosed and in metastatic breast cancer. Although a histopathologic assessment is an established surrogate marker for a treatment response, approximately 13%–25% of patients showing a pCR develop a systemic recurrence after 5 y of follow-up and 60%–70% of patients not showing a pCR remain free of a recurrence. The strength of a histopathologic assessment is the identification of a pCR after the completion of systemic treatment, whereas the strength of $^{18}$F-FDG PET is the identification of nonresponders early during treatment.

The lack of clinical acceptance of $^{18}$F-FDG PET–based treatment stratification is driven by several factors. No SUV cutoffs for separating responders and nonresponders have been established or validated by patient outcome data. Baseline tumor metabolic activities and subsequent treatment-induced changes are different for distinct molecular subtypes, and specific SUV cutoffs are necessary for different breast cancer subtypes. Neoadjuvant treatment is generally based on aggressive chemotherapy regimens, and alternative treatment approaches—other than immediate surgery—are often not available. An important contribution of $^{18}$F-FDG PET could be directing the addition of platinum-based chemotherapy in poorly responding triple-negative breast cancer patients, but such an approach requires prospective validation. The possibility that tumor shrinkage in patients not achieving a pCR might enable breast-conserving surgery is important. The lack of standardization of clinical $^{18}$F-FDG PET imaging outside research protocols leads to high variations in $^{18}$F-FDG uptake in tissue and hinders clinical acceptance. Although less applicable for routine clinical use because of cost and availability, PET/MRI would allow for a direct comparison between changes in glucose metabolism and the accumulation and kinetic behavior of MRI contrast agents, and diffusion-weighted MRI could be used to examine changes in water restriction in the same patients. Such comparisons would be helpful for identifying the strengths and weaknesses of $^{18}$F-FDG PET versus or combined with MRI.

In the metastatic setting, the potential clinical application of $^{18}$F-FDG PET is based on no significant changes in tumor metabolic activity, indicating ineffective treatment. More alternative treatment options are available in this setting, and treatment could be adapted earlier than with conventional imaging. The number of prospective trials is still too small to support this approach in clinical practice.

Given the high cost of applying ineffective treatments in (breast) cancer patients and the cost of evaluating PET imaging, alternative national and international funding methods for performing the prospective clinical multicenter trials needed to establish PET treatment monitoring must be found.

DISCLOSURE

This publication was made possible by the Clinical and Translational Science Collaborative of Cleveland, by UL1TR000439 and KL2TR000440 from the National Center for Advancing Translational Sciences (NCATS) component of the National Institutes of Health and NIH Roadmap for Medical Research, by Case Western Reserve University, and by Ohio Third Frontier Funding. No other potential conflict of interest relevant to this article was reported.

REFERENCES


18F-FDG PET/CT for Monitoring of Treatment Response in Breast Cancer

Stefanie Avril, Raymond F. Muzic, Jr., Donna Plecha, Bryan J. Traughber, Shaveta Vinayak and Norbert Avril

Doi: 10.2967/jnumed.115.157875

This article and updated information are available at:
http://jnm.snmjournals.org/content/57/Supplement_1/34S

Information about reproducing figures, tables, or other portions of this article can be found online at:
http://jnm.snmjournals.org/site/misc/permission.xhtml

Information about subscriptions to JNM can be found at:
http://jnm.snmjournals.org/site/subscriptions/online.xhtml