Whole-Body $^{18}$F-FDG PET and Conventional Imaging for Predicting Outcome in Previously Treated Breast Cancer Patients

Duska Vranjesevic, MD; Jean Emmanuel Filmont, MD; Joubin Meta, MS; Daniel H. Silverman, MD, PhD; Michael E. Phelps, PhD; Jyotsna Rao, MD; Peter E. Valk, MD; and Johannes Czernin, MD

1 Department of Nuclear Medicine, UCLA School of Medicine, Los Angeles, California; and 2 Northern California PET Imaging Center, Sacramento, California

This study was conducted to determine the ability of $^{18}$F-FDG PET and conventional imaging (CI) to predict the outcomes in breast cancer patients who have previously undergone primary treatment. Methods: The study population consisted of 61 female patients (median age, 54 y; range, 32–91 y) who were reevaluated with $^{18}$F-FDG PET and CI after treatment. The median interval between the last treatment and PET was 0.4 y (range, 0–16 y). PET was performed within 3 mo of CI (median interval, 25 d; range, 2–84 d). To determine the independent impact of PET on outcome, PET images were reinterpreted in a blind fashion. Availability of clinical information after PET scanning (21 of 110) was required for study inclusion. Study endpoints were clinical evidence of progression of disease or death. Results: Of 61 patients, 19 (31.1%) had no clinical evidence and 38 (62.3%) had evidence of residual or recurrent disease by the end of follow-up. Four patients (6.6%) had died. The positive and negative predictive values (PPV and NPV, respectively) of PET were 93% and 84%, respectively. CI yielded a PPV of 85% and an NPV of 59%. The prognostic accuracy of single whole-body PET was superior to that of multiple procedures with CI (90% vs. 75%; $P < 0.05$). Kaplan–Meier estimates of disease-free survival in patients with negative PET findings compared with those with positive PET findings revealed a significant difference between the 2 curves (log-rank test = 0.001). Kaplan–Meier estimates of disease-free survival stratified by CI results showed a marginally significant difference between CI-positive and CI-negative patients (log-rank test = 0.04). Conclusion: FDG PET can be used to improve prediction of the clinical outcome of previously treated breast cancer patients relative to what is achievable through CI alone.

Key Words: breast cancer; $^{18}$F-FDG PET; patient outcome; conventional imaging

Inclusion Criteria

The following inclusion criteria were required for the patient’s enrollment in the study: (a) availability of CI reports performed within 3 mo of PET, (b) availability of follow-up data for a minimum of 6 mo after scanning, and (c) unequivocal determination of clinical status at the time of the last clinical follow-up.

Clinical Scenario

After initial diagnosis, patients underwent a variety of treatments. Surgery (lumpectomy or mastectomy) was performed on all patients. Forty-six patients (75.4%) underwent chemotherapy and 34 (55.7%) received radiation therapy. The median interval between the last treatment and PET was 0.4 y (range, 0–16 y).

Clinical whole-body PET imaging was ordered for the following indications: Forty-two patients (68.8%) were being evaluated for residual or recurrent disease without any specified reason, 10 (16.4%) had increasing levels of serum tumor markers, and 9 (14.8%) had equivocal or suspicious findings by CI. The follow-up period after PET averaged 21 ± 12 mo (median, 17 mo; range, 6–42 mo).

The clinical status, as determined by the treating physician or chart review (or both), at the time of the last follow-up was used to categorize patients as free of disease or having disease. The latter group consisted of patients who at the time of last follow-up had stable or progressive disease or had died.

Clinical PET Imaging Protocol

Patients fasted for at least 6 h before receiving an intravenous injection of 370–555 MBq 18F-FDG. PET was performed with an ECAT EXACT or HR+ system (CTI/Siemens, Knoxville, TN).

The standard clinical imaging protocol started 45–60 min after tracer injection. Images were acquired over 6–8 bed positions. Images were acquired with (n = 11) or without (n = 50) attenuation correction and were reconstructed using filtered backprojection or algebraic algorithms (23).

Image Interpretation and Classification

To determine the independent prognostic value of PET, images were reinterpreted by an experienced reader who was unaware of other clinical findings. Imaging findings were dichotomized as positive or negative for residual or recurrent cancer.

Written clinical reports of conventional images were reviewed and classified as (a) negative, if all imaging tests were negative for disease; (b) equivocal, when abnormal findings were present on any imaging test but were not interpreted as suspicious for malignancy; (c) suspicious, if any test result was clearly described as suspicious for malignancy; or (d) positive, if findings were described as consistent with malignancy. To dichotomize the data, negative and equivocal findings were subclassified as negative, and suspicious and positive findings were categorized as positive.

The clinical impression was derived from follow-up CI studies on 56 patients, biopsy of abnormalities on 18 patients, and rising tumor markers in 3 patients. The outcome of patients with true-positive PET but false-negative CI findings was verified by CI follow-up in all 6 patients, rising levels of tumor markers in 1 patient, and biopsy in another 2 patients. This information was then evaluated and categorized as stable or progressive disease. Thus, patients were considered stable if the above tests did not indicate progression of disease.

Statistical Analysis

Positive and negative predictive values (PPV and NPV, respectively) and accuracy were calculated by standard methods. Categoric data were examined using the $\chi^2$ test, with Fisher’s exact test if expected cell counts were <5. Probability of disease-free survival was computed by Kaplan–Meier analysis. The log-rank test was used to evaluate the differences between Kaplan–Meier curves. All statistical tests were performed at the 5% level of significance.

RESULTS

Reevaluation by PET after treatment was performed within 3 mo of CI (median, 25 d; range, 2–84 d). A total of 189 CI procedures, or 3.2 ± 1.4 procedures per patient, was performed. The imaging procedures performed included whole-body bone scanning (n = 40), chest CT (n = 35), abdominopelvic CT (n = 33), chest radiography (n = 24), MRI of the breast and other MRI (n = 26), sonography of the breast (n = 10), mammography (n = 16), and other procedures (n = 5).

At the time of the last follow-up, 19 of 61 patients (31.1%) were considered free of disease, whereas 38 of 61 patients (62.3%) had clinical evidence of stable or progressive disease. The remaining 4 patients (6.6%) had died.

CI yielded negative results in 22 patients (36%) (negative findings, n = 14; equivocal findings, n = 8) and positive results in 39 patients (64%) (positive findings, n = 24; suspicious findings, n = 15). PET showed negative findings in 19 patients (31.1%) patients, whereas 42 patients (68.9%) had evidence for disease on PET.

PET and CI were discordant in 46 of 61 patients (75%), 33 in which they were both positive and 13 in which they were both negative. All 33 of these positive cases were true-positive cases (PPV = 100%), and 10 of these negative cases were true-negative cases (NPV = 77%). PET and CI were discordant in the remaining 25% of patients, 9 which were positive on PET but negative on CI and 6 which were negative on PET but positive on CI. PET correctly predicted the clinical outcome in 12 of these 15 discordant cases (80%), whereas CI was correct in only 3 cases (20%), all of which were positive on PET but negative on CI (Table 1).

Using the clinical outcome as the gold standard, blind PET evaluation yielded a sensitivity of 93% versus 79% for CI (P < 0.05). The specificity tended to be higher for PET than for CI (84% vs. 68%; P = 0.065). The PPV of PET was similar to that of CI (93% vs. 85%; P = 0.24). The NPV and predictive accuracy of blind PET evaluation were significantly higher than those of CI (84% vs. 59% [P < 0.05] and 90% vs. 75% [P < 0.05]) (Table 2).

TABLE 1

<table>
<thead>
<tr>
<th>Imaging Findings Versus Clinical Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Negative (n = 22)</td>
</tr>
<tr>
<td>Positive (n = 39)</td>
</tr>
<tr>
<td>Negative (n = 6)</td>
</tr>
</tbody>
</table>

326  THE JOURNAL OF NUCLEAR MEDICINE • Vol. 43 • No. 3 • March 2002
Kaplan–Meier estimates of disease-free survival in 19 patients with negative findings on PET were compared with those of 42 patients who had positive results on PET: The 2 survival curves differed significantly (log-rank test = 0.001). Kaplan–Meier estimates of disease-free survival in patients with negative CI findings (n = 22) compared with those with positive findings (n = 39) revealed a marginally significant difference between the 2 curves (log-rank test = 0.04) (Fig. 1).

DISCUSSION

This study shows the higher accuracy of 18F-FDG PET compared with combined CI modalities for predicting outcome in breast cancer patients who were reevaluated after primary treatment. Breast cancer patients commonly undergo numerous imaging studies after initial treatment even though the therapeutic options for patients with recurrent or residual disease are limited. Nevertheless, detection of local recurrence and metastatic disease can impact therapy. Local recurrence and axillary lymph node involvement may be treated surgically or with radiation therapy, whereas mediastinal lymph node involvement or distant metastases typically require chemotherapy or radiation therapy (or both) (24). The most frequently used staging examinations include plain roentgenography of the chest and bones; CT of the chest, abdomen, and pelvis; sonography; MRI; mammography; and whole-body radionuclide bone scanning.

In this study, patients underwent an average of 3.2 CI tests within 3 mo of PET and many more throughout the course of the disease. This study indicates that PET provides more accurate prognostic information than the combination of other imaging tests.

All patients enrolled in the study underwent reevaluation after treatment by CI and were followed for almost 2 y. Most patients (42/61 [68.9%]) had clinical evidence for disease at the time of the last follow-up. Clinical outcome was clearly defined in all patients and was used as a gold standard for determining the predictive accuracy of PET and CI. The predictive accuracy of PET was significantly better than that of CI (90% vs. 75%; P < 0.05). However, false-positive and false-negative findings on PET were found in 6 of 61 patients.

In all 3 cases with false-negative findings, rising levels of serum tumor markers represented the indication for PET. All 3 of these patients also had negative findings on CI. Interestingly, 2 of 3 patients with false-negative results had lobular carcinoma. Lobular carcinoma accounts for only 7%–10% of all invasive breast cancers (25–27) and exhibits lower glucose metabolic activity than invasive ductal carcinoma (28,29). These 3 patients remained asymptomatic for a mean period of 14.5 mo after PET imaging, at which time the metastatic disease was discovered. The sites of the metastatic lesions that were identified eventually were liver

<table>
<thead>
<tr>
<th>Parameter</th>
<th>CI</th>
<th>Blind PET interpretation</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%)</td>
<td>78.6 ± 10.3</td>
<td>92.9 ± 6.5</td>
<td>*</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>68.4 ± 11.7</td>
<td>84.2 ± 9.2</td>
<td></td>
</tr>
<tr>
<td>PPV (%)</td>
<td>84.6 ± 9.1</td>
<td>92.9 ± 6.5</td>
<td></td>
</tr>
<tr>
<td>NPV (%)</td>
<td>59.1 ± 12.3</td>
<td>84.2 ± 9.2</td>
<td>*</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>75.4 ± 10.8</td>
<td>90.2 ± 7.5</td>
<td>*</td>
</tr>
</tbody>
</table>

*P < 0.05 for PET vs. CI.

FIGURE 1. (A) Kaplan–Meier estimate of disease-free survival in 42 patients with positive PET findings compared with 19 patients with negative PET findings. P = 0.001 (log-rank test). (B) Kaplan–Meier estimate of disease-free survival in 22 patients with negative CI findings compared with 39 patients with positive CI findings. P = 0.04 (log-rank test).
and bone in 1 patient, bone only in a second patient, and the anterior abdominal wall in the third patient.

PET yielded a significantly higher sensitivity (92.2%) than CI (78.6%) attributed to the larger number of false-negative results on CI. In 6 CI-negative cases, PET revealed recurrence, which was local in 1 patient, involved axillary and supraclavicular nodes in 1 patient (Fig. 2), involved mediastinal lymph nodes in 2 patients, and involved bone metastases in 2 patients. These results support previous studies that showed that lymph node metastasis and bone metastases are detected significantly more often on FDG PET than on CT or MRI (24). The CI tests performed on these patients included bone scanning (n = 5), CT or MRI of the chest (n = 6) and of the abdomen and pelvis (n = 4), and sonography (n = 1). All patients with positive findings on PET but negative findings on CI were treated with chemotherapy (n = 6), radiation (n = 3), and other treatments (n = 1).

PET yielded false-positive results in 3 patients. The false-positive results by PET resulted from uptake in a benign pleural effusion in 1 patient, physiologic muscle activity that was misinterpreted as lymph node involvement in a second patient, and uptake attributed to mastitis in the third patient.

In the group with positive CI findings (n = 39), PET was concordantly positive in 33 cases, and all of these patients had clinical evidence of relapse. The remaining 6 patients had false-positive results by CI. In 2 of these patients, CT showed suspicious lesions in the liver, conventional bone scan findings were suspicious for metastatic disease in 3 patients, and in 1 patient local recurrence was suspected by MRI of the breast. These 6 patients were monitored by other imaging tests and remained free of disease at the time of the last follow-up. Overall, among all cases in which PET and CI were discordant, PET was correct 4 times more frequently than CI (12/61 vs. 3/61 cases) and was correct in all discordant cases in which PET findings were negative.

This study has several limitations. First, patients were studied retrospectively, and only 25% of the 235 patients who underwent PET imaging after treatment for breast cancer had adequate follow-up. Specifically, no consistent information was available to further elucidate the exact time and location of disease recurrence in patients who had positive PET findings but negative CI findings. Nevertheless, to date, this is the largest study to determine the predictive value of 18F-FDG PET for patient outcome in treated breast cancer patients.

Second, the disease-free survival of breast cancer patients depends on accurate staging of the disease and also on further therapy initiated because of positive PET findings. In this study, 6 patients had true-positive PET findings but false-negative CI findings. All of these patients underwent further chemotherapy (n = 6) or radiation therapy (n =3) and surgery (n = 1) (or all 3 treatments). Note that despite these therapeutic interventions, all patients had evidence of disease at the time of follow-up (3 with stable disease, 1 with progressive disease, and 2 who had died). Thus, 18F-FDG PET affected the management of these patients. However, the effect of these interventions cannot be determined with certainty from the current data. Treatment did not result in remission but might have delayed the progression of disease in some patients.

Third, dichotomizing the CI data as positive (for suspicious and positive findings) and negative (for negative and equivocal findings) might have skewed the results toward favoring PET. To address this issue, a subanalysis was performed that included only those patients who had unequivocal positive or negative findings on CI (n = 38). The sensitivity (81%) and the PPV (87.5%) of CI were similar to those of the entire population. The specificity and the NPV were 75% and 64%, respectively. The overall predictive accuracy of CI for this subset of the population was 79%, significantly lower compared with the predictive accuracy of PET for the same subset of the population (92%).

Fourth, written reports of CI tests, rather than rereading of studies, were used to determine the conventional test result. This approach was chosen because of the practical difficulties of obtaining original images from numerous different referral sites. On the other hand, PET images were reread because initial PET reading had been generally performed with full knowledge of the patient’s history and correlative image findings. This approach might have placed PET at an unfair advantage if we had not reread the PET scans. However, the separate analysis based on written clinical PET reports showed a significantly higher sensitiv-
ity and NPV of PET than those of CI findings, as with the blind readings, and the overall predictive accuracy of non-blind PET findings did not differ significantly from that of the blind PET interpretation.

Fifth, most PET images were acquired without attenuation correction because attenuation correction and iterative image reconstruction were not routinely performed in our clinic before January 2000. However, the attenuation correction is unlikely to change the diagnostic or predictive accuracy of PET (30,31).

Finally, study inclusion required that PET and CI needed to be performed within 3 mo of each other. Because CI was performed before PET, disease might have recurred during this time interval. To evaluate the effect of the study design on our results, we narrowed the time interval between CI and PET to 6 wk (median, 15 d; range, 2–40 d). Thirty-two patients underwent evaluation by PET and CI within that interval. The PPV, NPV, and predictive accuracy of PET were 95%, 90%, and 94%, respectively, which was significantly higher than the corresponding values for CI (75%, 50%, and 69%, respectively; P < 0.05). Moreover, the most significant impact of PET overall was an increase in the cases with true-negative PET but false-positive CI. Thus, the length of the time interval between PET and conventional studies was unlikely to have accounted for the current results.

CONCLUSION

18F-FDG PET can be used to improve prediction of the clinical outcome of previously treated breast cancer patients relative to what is achievable through CI alone.

REFERENCES