Determinants of Diagnostic Performance Of [F-18]Fluorodeoxyglucose Positron Emission Tomography for Axillary Staging in Breast Cancer

Jacobus J. M. van der Hoeven, MD,*† Otto S. Hoekstra, PhD,†‡ Emile F. I. Comans, PhD,† Rik Pijpers, PhD,† Robert P. A. Boom, PhD,§ Dick van Geldere, PhD,§ Sybren Meijer, PhD,¶ Adriaan A. Lammertsma, PhD,† and Gerrit J. J. Teule, PhD†

*From the Departments of *Internal Medicine and §Surgery, Ziekenhuis Amstelveen, Amstelveen, The Netherlands, and the Departments of †Nuclear Medicine, ‡Clinical Epidemiology & Biostatistics, and ¶Surgery, VU University Medical Center, Amsterdam, The Netherlands*

Objective

To prospectively investigate determinants of the accuracy of staging axillary lymph nodes in breast cancer using [F-18]fluorodeoxyglucose positron emission tomography (FDG PET).

Methods

Patients with primary operable breast cancer underwent FDG PET of the chest followed by sentinel node biopsy (SNB, $n =$ 47) and/or complete axillary lymph node dissection (ALND, $n = 23$). PET scans were independently interpreted by three observers in a blinded fashion with respect to the FDG avidity of the primary tumor and the axillary status. The results were compared to histopathological analyses of the axillary lymph nodes. Clinicians were blinded to the PET results.

Results

Axillary lymph node specimens and FDG PET scans were evaluated in 70 patients (59% cT1). Overall, 32 (46%) had lymph node metastases as established by SNB (18/47) or

Several studies have investigated [F-18]fluorodeoxyglucose positron emission tomography (FDG PET) for staging lymph nodes in breast cancer. Reported sensitivities range from 50% to 100% and specificities from 66% to 100% .¹⁻⁸ There is no straightforward explanation for these variable

Correspondence: JJM van der Hoeven, MD, Department of Internal Medicine, Ziekenhuis Amstelveen, Laan van de Helende Meesters 8, 1186 AM, Amstelveen, The Netherlands.

E-mail: jjho@zha.nl Accepted for publication February 19, 2002.

DOI: 10.1097/01.SLA.0000032954.55626.4E

ALND (14/23), 20 of which were confined to a single node. The overall sensitivity of FDG PET was 25%, with a specificity of 97%. PET results were false-negative in all 18 positive SNBs and true-positive in 8/14 in the ALND group. The performance of FDG PET depended on the axillary tumor load and the FDG avidity of the primary tumor. Intense uptake in the primary tumor was found in only 57% of the patients, and this was independent of the size. There was excellent interobserver agreement of visual assessment of FDG uptake in primary tumor and axillary lymph nodes.

Conclusions

The sensitivity of FDG PET to detect occult axillary metastases in operable breast cancer was low, and it was a function of axillary tumor load and FDG avidity of the primary tumor. Even though the clinical relevance of occult disease detected by SNB needs to be confirmed, it is suggested that FDG PET in these patients should be focused on exploiting its nearly perfect specificity and the potential prognostic relevance of variable FDG uptake.

results. They may relate to different implementations of the gold standard (histopathological examination of surgical specimens), to differences in patient populations, or to differences in PET scan procedures or image assessment criteria.⁹ Some studies have claimed a very high sensitivity of FDG PET in axillary staging, $2,5,7,8$ to the extent that the negative predictive values of PET and the sentinel node biopsy (SNB) appear to be similar.⁵ Since in the published literature PET was typically compared to the results of histopathological examinations of axillary lymph node dissections (ALNDs) and not SNB, this counterintuitive suggestion needs to be investigated in a direct comparison.

Figure 1. Study protocol.

Further, the FDG avidity of malignant breast tumors is variable and relatively lower than in, for instance, lung cancer, 9,10 but it is unclear whether and how this affects the reproducibility of image interpretation and the accuracy of FDG PET in axillary staging.

The purpose of this prospective study was to investigate the performance of FDG PET as a function of several of these determinants: the axillary status as defined with SNB and immunohistochemistry, the FDG avidity of the primary tumor, and PET scan reading criteria.

METHODS

Patient Population

Patients from the VU University Medical Center and the community hospital Amstelveen with cytologically or histologically proven and operable breast cancer, newly presenting between April 1997 and April 2000, were eligible for the study. Patients had to give informed consent; patients with diabetes were excluded.

Clinical Procedures

Axillary staging was performed according to standard hospital procedures (Fig. 1). During the first phase of the study, complete ALND was performed routinely as part of SNB validation and implementation studies.¹¹ SNBs were performed using the combined blue dye/radiocolloid approach.12 At a later stage, ALND was performed only after a tumor-positive SNB and in patients not eligible for the SNB procedure according to national guidelines.¹³ After a negative SNB or ALND, patients were followed clinically for at least 18 months. Surgeons and pathologists were blinded to the PET results.

Pathology

If one slice of the sentinel node did not reveal metastasis, slices at four to six other levels were made from this node and examined for the presence of metastases with HE and CAM5.2. For ALND specimens, slices at one or two levels (for nodes ≤ 1 cm and ≥ 1 cm, respectively) were stained with CAM 5.2 and examined for the presence of metastases.

FDG PET

FDG PET scans were performed within 1 week before surgery. Before the PET scan, patients fasted for at least 6 hours. Serum glucose was measured just before the intravenous administration of 370 MBq FDG in the arm opposite to the site of the primary breast tumor (within normal limits in all subjects). After a resting period of 60 minutes, patients were transferred to the scanning room. Emission scans of the chest (2-D mode, two bed positions of 15 minutes) were acquired, with the patient in the supine position, using a dedicated PET scanner (ECAT EXACT HR $+$, CTI/ Siemens). PET scans were corrected for decay, scatter, and randoms and reconstructed as 128×128 matrices using filtered back-projection with a Hanning filter (cutoff 0.5 cycles/pixel), resulting in a transaxial spatial resolution of ~7 mm full-width at half-maximum.

PET images were visually analyzed by three independent observers blinded to all clinical data. They scored focally enhanced FDG uptake in primary tumor and axillary nodes on a scale of 0 to 3 for absent, faint, moderate, or intense uptake relative to normal tissue, respectively. The score assigned by at least two observers was used as final sore. This allowed classification of all cases.

Statistical Analysis

Interrater agreement was assessed with the intraclass correlation coefficient (ICC); values range from 0 to 1. Values close to 0 indicate poor agreement between observations (ie, most of the variance is due to measurement error). Values close to 1 indicate high agreement. The Fisher exact test was used to identify differences in proportions and the Cochrane Armitage trend test for trend analyses of proportions.

RESULTS

During the study period, 70 patients were enrolled. The characteristics of these patients, their clinical stage, and the histologic type of the primary tumors are listed in Table 1. The majority presented with a small primary tumor (59% $\leq cT1$) and a clinically unremarkable axilla (71%). The diagnosis had been established before the PET scan with fine-needle aspiration in 58 patients and with excisional biopsy in the other 12 patients (in the latter group 15–83 days [mean 28] before the scan). In the latter group of patients, variable residual FDG uptake at the original site of

Table 1. PATIENT CHARACTERISTICS

	Excisional Biopsy	No Excisional Biopsy	Total
Number of patients	12	58	70
Age mean (sd)	51(9)	60 (13)	58 (13)
Clinical T stage			
T ₀	1	3	4 (6%)
T1	7	30	37 (53%)
T ₂	1	17	18 (26%)
T ₃		4	4(6%)
T ₄		3	3(4%)
Unknown	3	1	4(6%)
SN biopsy	12	35	47
pT size (mm), mean (sd)	16(7)	25(12)	23(11)
pΝ			
N ₀	7	31	38 (54%)
N ₁	5	27	32 (46%)
Type of tumor			
Ductal	9	40	49 (70%)
Lobular	1	8	9(13%)
Other	$\overline{2}$	10	12 (17%)

the primary tumor was observed in all but one patient (faint, $n = 6$; moderate, $n = 4$; intense, $n = 1$). In nine cases, no residual malignancy was found in the surgical breast specimen. In the three remaining patients with residual tumor after excisional biopsy, the PET scan showed absent, faint, and intense uptake, respectively.

No FDG uptake in the primary tumor was observed in 5 of the 58 patients. At pathologic examination, these tumors measured 10, 15, 19, 25, and 55 mm and were histologically classified as mucinous adenocarcinoma $(n = 1)$, ductal adenocarcinoma ($n = 2$), and infiltrating lobular carcinoma

 $(n = 2)$. In the remaining cases, FDG uptake was scored as intense in 57% (n = 33), moderate in 28% (n = 16), and faint in 7% $(n = 4)$. The interrater agreement of this classification proved to be excellent (ICC = 0.93 , 95% CI 0.89–0.96). Tumor size as defined by pathologic examination and FDG uptake were not associated ($n = 58$, $P = .31$; Cochrane Armitage trend test, $n = 58$).

Histologic examination revealed tumor-positive lymph nodes in 32 patients (46%); spread was confined to a single node in 19 patients. Eighteen of the 47 SNBs were tumorpositive (38%), with additional positive nodes in the ALND specimen in only 5 patients. In the routine ALND group, 14 of 23 had tumor-positive lymph nodes (61%), 8 of which were multiple.

For the entire group of 70 patients, the sensitivity of PET for axillary staging was low for any of the applied PET reading criteria (Table 2). PET scoring of the axilla showed a very low interobserver variability (ICC $= 0.97, 95\%$ CI 0.95–0.98). "At least moderately enhanced" axillary uptake appeared to have the best operating characteristics, with a sensitivity of 25% (95% CI 12–43) and a specificity of 97% (95% CI 86–100%). One patient had a positive PET scan (Fig. 2) and a negative SNB, and no additional ALND was performed. Since the primary tumor was estrogen receptorpositive, the patient received adjuvant therapy with tamoxifen. Two years after surgery, no recurrent disease has been detected. For the purpose of this study, the PET scan result was classified as a false-positive.

The sensitivity of PET was positively related to the axillary tumor load: PET detected only 2 of 20 axillas with single-node involvement versus 6 of 12 with multiple tumor-positive nodes ($P = .03$, Fisher exact test, see Table 2). An example of a positive PET scan in a patient with mul-

 $*$ PET considered positive if \geq 2 observers read intense uptake (3), moderate (2), faint (1), respectively.

† PET scored positive in case of at least moderately enhanced focal axillary FDG uptake.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Figure 2. FDG PET scan of a patient showing intense uptake by the primary tumor and moderate uptake in an axillary lymph node. Since the sentinel node was tumor negative, no additional ALND was performed. An example of a false-negative sentinel lymph node procedure?

tiple axillary lymph node metastases is shown in Figure 3. Further, the ability to detect these multiply involved nodes appeared to be associated with the FDG avidity of the primary tumor (6/11 for all 58 primary tumors, 6/9 in the 49 moderate/intense group, and 6/6 in the 33 primary tumors with intense FDG uptake; Cochrane Armitage Trend test, $P = .05$, one-sided).

In the subgroup of 49 patients with "adequate" (ie, at least moderately enhanced) FDG uptake in the primary tumor, the sensitivity of PET was 35% (95% CI 16–57%), and again was related to the axillary tumor load: PET detected 2 of 14 cases with single-node involvement versus 6 of 9 with multiple affected nodes ($P = .02$, Fisher exact test). FDG avidity of the primary tumor was not associated with the histopathological N status. In the SNB group with "adequate" FDG uptake in the primary tumor $(n = 29)$, 10 patients had a tumor-positive lymph node biopsy, but none had been detected by PET.

DISCUSSION

In the present study, FDG PET was not an accurate indicator of occult axillary metastases in breast cancer as established with SNB and immunohistochemistry. Further, the performance of FDG PET in axillary staging proved to be a function of the FDG avidity of the primary process and of the axillary tumor load.

Most FDG PET studies published so far have reported a high sensitivity for axillary lymph node involvement in breast cancer (Table 3). Greco et al even suggested that SNB or ALND may be avoided in case of a negative PET scan of the axilla.⁵ Our data do not support this suggestion: PET failed to detect any of the solitary tumor-positive

Figure 3. FDG PET scan of a patient showing intense uptake in both the primary tumor and axillary lymph nodes. ALND specimen revealed multiple lymph node metastases.

Table 3. DIAGNOSTIC PERFORMANCE OF FDG PET IN AXILLARY STAGING

Studies with at least 35 breast cancer patients published in the English literature until June 2001.

¹ Number of breast cancers included.

² Between parentheses: (mean; range).

³ Number of sections per lymph node; HE: hematoxylin-eosin, IHC: immunohistochemistry.

4 n.a.: not available from the manuscript.

⁵ Only data on the pN1 group.

sentinel nodes. The PET studies have included a generally more advanced spectrum of disease as compared with the SNB literature, which typically focuses on smaller tumors (T1) and clinically unremarkable axillas. Indeed, a clearly lower sensitivity has been reported for FDG PET in the T1 subsets $1,14$ compared to that of the overall patient sample. In our study, patients were stratified for SNB or ALND largely depending on their tumor size. As expected, a higher percentage of axillary involvement was found in the ALND group. Even though conceptually PET may very well detect very small lesions, the present evidence suggests that for peripheral lymph node staging, its sensitivity rapidly declines below 5 mm, even in tumors that are very avid for FDG.15,16 In our experience, about 50% of the positive SNBs contain less than 3 mm of tumor tissue.¹⁷ Therefore, the most likely explanation of the apparently lower sensitivity in our hands is the higher detection rate of lymph node metastases by the SNB procedure and the use of immunohistochemical analysis. Most PET studies have used H&E staining of ALND specimens as the gold standard. It has been shown that SNB analysis using immunohistochemistry and step-sectioning will upstage 30% of the patients staged as node negative by ALND and H&E staining.¹⁸ Indeed, with this technique, Greco et al⁵ found 23% pN1 in the T1 patients versus 33% in our study.

Microscopic disease will remain beyond the realm of imaging methods like PET, MRI, and ultrasound. In lowprevalence conditions, the SNB is clearly superior to exclude metastatic axillary involvement to avoid unnecessary surgery because of its near-perfect specificity. However, noninvasive techniques may serve to expand the SNB procedure toward larger tumors, where SNB experience is limited.¹⁹ If macroscopic tumor is present to the extent that lymphatic drainage is diverted away from the sentinel node,

preoperative imaging might serve to select the patients at risk of a false-negative SNB. Alternatively, the very high specificity of FDG PET may stratify patients in trials comparing neoadjuvant versus adjuvant therapy in node-positive patients. This is especially attractive if the level of FDG uptake indeed bears independent prognostic information.^{20,21} Finally, it is unclear in which patients removal of only a tumor-positive sentinel node might be a safe procedure. The role of adjuvant local therapy in such cases is debated, since most patients with a positive SNB will receive adjuvant treatment anyway, and since in approximately 60% the sentinel node appears to be the single positive node at complementary ALND.¹¹ Even though our sample was very small, one might speculate that in primary tumors very avid for FDG, a negative FDG PET scan of the axilla could justify the approach to resect only the sentinel node. The clinical relevance of minimal nodal disease can be studied only in a randomized trial comparing long-term outcomes of patient management guided by SNB or PET. This will confirm whether ALND can be safely avoided in patients with a negative FDG PET scan, $⁶$ and whether the</sup> microscopic amounts of tumor found by the SNB always justify adjuvant local or systemic therapy. 22

We used non-attenuation-corrected images in our study, since it has not been established that such correction improves the performance of PET in this setting.²³ As a consequence, we used a visual assessment of FDG uptake (rather than standardized uptake values), which proved to be a reproducible method. It could be argued that the true level of FDG uptake will be underestimated in small tumors (less than twice the resolution). In the present patient population, nine patients had tumors smaller than 14 mm, two of them smaller than 1 cm. However, since only one had a tumorpositive axilla, and seven had moderate or intense uptake in

the primary tumors, this cannot have affected the observed relation between FDG avidity and yield of PET in staging the axilla.

It appears from our data that there is a positive correlation between FDG avidity and the detection rate of axillary metastases. The observed variability of FDG uptake by primary breast tumors appears to be in agreement with the findings of Avril et al, 9 who reported a sensitivity of 64% to 80% for detection of primary breast cancer, depending on scoring criteria and tumor size. Even though the association between FDG uptake in the primary tumor and the detection rate of axillary metastases is biologically plausible, this is, to our knowledge, the first study to demonstrate it. Future studies may consider stratifying patients with respect to the FDG avidity of their tumors. For clinical practice, it remains to be shown whether biopsy specimens can be used to restrict PET to patients who are likely to have an FDG-avid t umor.^{21,24} Until then, patients who underwent an excisional biopsy of the primary tumor should not be included in FDG PET studies investigating its role in axillary staging. First, they lack information about the FDG avidity of the primary tumor. Second, our study has shown that most tumor sites had residual FDG uptake several weeks after the excisional biopsy. Since the excision proved to have been radical in most (9/12) cases, this is probably caused by an inflammatory reaction. It is expected, however, that fine-needle aspiration, the primary diagnostic method in the remaining 58 patients, cannot elicit such reaction to the extent that visual assessment of the FDG avidity of the primary tumor will be biased. Since core biopsies have replaced excisional biopsy in many clinical settings, the effect of core biopsies on FDG uptake as a function of time needs to be investigated.

In conclusion, if the microscopic tumor load disclosed by the SNB is clinically relevant, FDG PET does not reliably exclude it, not even in patients with FDG-avid tumors, who are predicted to have an optimal detection rate of axillary metastases by FDG PET. Instead, the excellent and reproducible specificity of FDG PET, as well as the potential prognostic relevance of FDG uptake, suggests a role for PET in trials looking at preoperative chemotherapy in nodepositive patients.

References

- 1. Avril N, Dose J, Janicke F, et al. Assessment of axillary lymph node involvement in breast cancer patients with positron emission tomography using radiolabeled 2-(fluorine-18)-fluoro-2-deoxy-D-glucose. J Natl Cancer Inst 1996; 88:1204–1209.
- 2. Adler LP, Crowe JP, al Kaisi NK, et al. Evaluation of breast masses and axillary lymph nodes with [F-18] 2-deoxy-2-fluoro-D-glucose PET. Radiology 1993; 187:743–750.
- 3. Adler LP, Faulhaber PF, Schnur KC, et al. Axillary lymph node metastases: screening with [F-18]2-deoxy-2-fluoro-D-glucose (FDG) PET. Radiology 1997; 203:323–327.
- 4. Scheidhauer K, Scharl A, Pietrzyk U, et al. Qualitative [18F]FDG positron emission tomography in primary breast cancer: clinical relevance and practicability. Eur J Nucl Med 1996; 23:618–623.
- 5. Greco M, Crippa F, Agresti R, et al. Axillary lymph node staging in breast cancer by 2-fluoro-2-deoxy-D-glucose-positron emission tomography: clinical evaluation and alternative management. J Natl Cancer Inst 2001; 93:630–635.
- 6. Yutani K, Shiba E, Kusuoka H, et al. Comparison of FDG-PET with MIBI-SPECT in the detection of breast cancer and axillary lymph node metastasis. J Comput Assist Tomogr 2000; 24:274–280.
- 7. Smith IC, Ogston KN, Whitford P, et al. Staging of the axilla in breast cancer: accurate in vivo assessment using positron emission tomography with 2-(fluorine-18)-fluoro-2-deoxy-D-glucose. Ann Surg 1998; 228:220–227.
- 8. Utech CI, Young CS, Winter PF. Prospective evaluation of fluorine-18 fluorodeoxyglucose positron emission tomography in breast cancer for staging of the axilla related to surgery and immunocytochemistry. Eur J Nucl Med 1996; 23:1588–1593.
- 9. Avril N, Rose CA, Schelling M, et al. Breast imaging with positron emission tomography and fluorine-18 fluorodeoxyglucose: use and limitations. J Clin Oncol 2000; 18:3495–3502.
- 10. Torizuka T, Zasadny KR, Recker B, et al. Untreated primary lung and breast cancers: correlation between F-18 FDG kinetic rate constants and findings of in vitro studies. Radiology 1998; 207:767–774.
- 11. Borgstein P, Pijpers R, Comans EF, et al. Sentinel lymph node biopsy in breast cancer: guidelines and pitfalls of lymphoscintigraphy and gamma probe detection. J Am Coll Surg 1998; 186:275–283.
- 12. Pijpers R, Meijer S, Hoekstra OS, et al. Impact of lymphoscintigraphy on sentinel node identification with technetium-99m-colloidal albumin in breast cancer. J Nucl Med 1997; 38:366–368.
- 13. Roumen RM, Pijpers HJ, Thunnissen FB, et al. Summary of the guideline: Dutch Work Group "Sentinel Node Biopsy for Breast Cancer." Ned Tijdschr v Geneesk 2000; 144:1864–1867.
- 14. Crippa F, Agresti R, Seregni E, et al. Prospective evaluation of Fluorine-18-FDG PET in presurgical staging of the axilla in breast cancer. J Nucl Med 1998; 39:4–8.
- 15. Wagner JD, Schauwecker DS, Davidson D, et al. FDG-PET sensitivity for melanoma lymph node metastases is dependent on tumor volume. J Surg Oncol 2001; 77:237–242.
- 16. Crippa F, Leutner M, Belli F, et al. Which kinds of lymph node metastases can FDG PET detect? A clinical study in melanoma. J Nucl Med 2000; 41:1491–1494.
- 17. Torrenga H, Licht J, van der Hoeven JJM, et al. Axillary lymph node staging in breast cancer by 2-fluoro-2-deoxy-d-glucose-positron emission tomography: clinical evaluation and alternative management [letter]. J Natl Cancer Inst 2000; 93(21):1659–60.
- 18. Torrenga H, Rahusen FD, Meijer S, et al. Sentinel node investigation in breast cancer: detailed analysis of the yield from step sectioning and immunohistochemistry. J Clin Pathol 2001; 54:550–552.
- 19. Bedrosian I, Reynolds C, Mick R, et al. Accuracy of sentinel lymph node biopsy in patients with large primary breast tumors. Cancer 2000; 88:2540–2545.
- 20. Oshida M, Uno K, Suzuki M, et al. Predicting the prognoses of breast carcinoma patients with positron emission tomography using 2-deoxy-2-fluoro[f-18]-d-glucose. Cancer 1998; 82:2227–2234.
- 21. Molthoff CFM, Bos R, van der Hoeven JJM, et al. Correlative in vitro and in vivo studies with FDG PET in breast cancer patients. J Nucl Med 2001; 42:26P.
- 22. van der Wall E. The sentinel node in breast cancer: implications for adjuvant treatment? Eur J Nucl Med 1999; 26:S17–S19.
- 23. Nakamoto Y, Chang AE, Zasadny KR, et al. Comparison of attenuation-corrected and non-corrected FDG-PET images for axillary nodal staging in newly diagnosed breast cancer. J Nucl Med 2001; 42:82P.
- 24. Avril N, Menzel M, Dose J, et al. Glucose metabolism of breast cancer assessed by 18F-FDG PET: histologic and immunohistochemical tissue analysis. J Nucl Med 2001; 42:9–16.