Differentiating recurrent tumor from radiation necrosis: time for re-evaluation of positron emission tomography?

P E Ricci, J P Karis, J E Heiserman, E K Fram, A N Bice and B P Drayer

http://www.ajnr.org/content/19/3/407
Differentiating Recurrent Tumor from Radiation Necrosis: Time for Re-evaluation of Positron Emission Tomography?

Peter E. Ricci, John P. Karis, Joseph E. Heiserman, Evan K. Fram, Alden N. Bice, and Burton P. Drayer

Our purpose was to evaluate the ability of FDG PET to differentiate recurrent tumor from posttherapy radiation necrosis.

METHODS: MR images, PET scans, and medical records of 84 consecutive patients with a history of a treated intracranial neoplasm were evaluated retrospectively. In all patients, recurrent tumor or radiation necrosis was suggested by clinical or MR findings. Metabolic activity of the PET abnormality was compared qualitatively with normal contralateral gray and white matter.

RESULTS: PET findings were confirmed histologically in 31 patients. With contralateral white matter as the standard of comparison, the PET scan sensitivity and specificity were found to be 86% and 22%, respectively. With contralateral gray matter as the reference standard, the sensitivity and specificity became 73% and 56%, respectively. Overall, nearly one third of the patients would have been treated inappropriately in either scheme had the PET scan been the sole determinant of therapy.

CONCLUSION: Our data suggest that the ability of FDG PET to differentiate recurrent tumor from radiation necrosis is limited. Both false-positive and false-negative PET scan results contributed to unacceptably low sensitivity and specificity values.

Over the past several decades, therapy for patients with primary and metastatic CNS tumors has become more aggressive as neurosurgeons, Neurooncologists, and radiation oncologists try to cure patients or at least provide them with a longer disease-free survival. Surgical resection and chemotherapy alone have proved to be insufficient therapy in many instances. As a result, the various forms of radiotherapy, including high-dose external beam radiation, radiosurgery, and radioactive seed implantation, have all become important therapeutic adjuncts. The end result is that radiation necrosis is being seen with increased frequency.

Unfortunately, distinguishing between radiation necrosis and recurrent or viable residual tumor has proved to be a particularly difficult task. Differentiation on the basis of clinical signs and symptomatology has not been possible. Similarly, computed tomography (CT) and magnetic resonance (MR) changes are not specific (1–5). On studies derived by those conventional imaging techniques, both radiation necrosis and recurrent tumor can exhibit an increased volume of enhancement, increased edema and mass effect, and regions of frank tissue necrosis or cavitation (1). This lack of specificity has prompted investigation into other imaging techniques with the hope of finding a more reliable clinical tool.

Early work with [18F] 2-fluoro-2-D-deoxyglucose (FDG) positron emission tomography (PET) was encouraging. A preliminary study by Patronas and DiChiro at the National Institutes of Health in the early 1980s suggested that FDG was both 100% sensitive and specific in its ability to differentiate recurrent tumor from radiation necrosis (6). Several subsequent studies supported the high sensitivity and specificity values reported by these investigators (7–12). However, more recent studies have questioned the efficacy and usefulness of FDG PET for this purpose, with specificities as low as 40% reported (13, 14).
Methods

We retrospectively reviewed medical records and FDG PET and MR studies of 84 consecutive patients with a history of a glial neoplasm treated with some combination of radiation, surgery, and/or chemotherapy. All patients had been examined clinically and radiologically over a 2-year period from May 1993 to June 1995. Each patient was thought to have recurrent tumor or radiation-induced necrosis on the basis of clinical symptoms or MR findings. Typical MR findings believed to be consistent with tumor recurrence included increased enhancement on T1-weighted images, an increase in the volume of T2 signal hyperintensity, necrosis and/or cavitation, and increased mass effect.

MR imaging was performed on a 1.5-T unit. Noncontrast and contrast-enhanced T1-weighted sagittal (700/16/1 [repetition time/echo time/excitations]) and axial (800/18/1) images were obtained along with noncontrast axial intermediate (2500/30/1) and T2-weighted (2500/90/1) images. Contrast-enhanced images were obtained immediately after intravenous administration of 0.1 mmol/kg of gadopentetate dimeglumine or gadoxamide.

PET scans were obtained on a GE 4096 PET scanner, which has an in-plane resolution of 6 mm, full width at half maximum (15). With patients in a resting state, 10 mCi of FDG was injected intravenously. After a 30-minute uptake period, patients were placed into the scanner. Head position and scan planes were oriented similarly to those used for MR imaging (approximately zero degrees to the orbitomeatal line) to facilitate comparison of the two studies. During the 25-minute imaging time, approximately 15 contiguous axial sections were obtained covering a 10.5-cm in-plane field of view. Filtered back projection with a calculated attenuation coefficient was used for image reconstruction. Images were reconstructed into a 128 × 128 matrix with a pixel size of 2 × 2 mm. Images were displayed in axial, sagittal, and coronal planes and were viewed on both hard copy and computer formats. The average time interval between PET and MR imaging was 8.5 days; the longest time interval between studies was 31 days.

Each of the PET and MR studies was interpreted by a CAQ-certified neuroradiologist at the time of the examination. PET scans were evaluated qualitatively by visual inspection and compared with the MR images to ensure that activity in the region of interest did not correspond to gray matter. The metabolic activity of each lesion was characterized as hypometabolic, isometabolic, or hypermetabolic relative to normal contralateral white matter. Lesions that were hypermetabolic relative to normal contralateral white matter were believed to be consistent with tumor recurrence; activity less than or equal to white matter was considered compatible with radiation necrosis. At the time this retrospective study was performed, a second neuroradiologist evaluated the PET and MR studies and graded lesion activity using a four-point grading system. A grade of 0 was assigned if there was no appreciable metabolic activity (eg, as seen in a surgical cavity). Grade 1 was assigned if activity was similar to normal contralateral white matter. Grade 2 if activity was between gray and white matter, and grade 3 if activity was greater than or equal to contralateral gray matter. Grade 0 and 1 lesions were interpreted as consistent with radiation necrosis; grade 2 and 3 lesions were considered to be consistent with recurrent tumor. There were no disagreements in interpretation between the two neuroradiologists. In the second part of the study, we attempted to improve the specificity of FDG PET for the diagnosis of recurrent tumor by comparing lesion activity with normal contralateral white matter. In this schema, only lesions isometabolic or hypermetabolic relative to gray matter (grade 3) were considered compatible with tumor recurrence.

Of the 84 patients whose PET scans were reviewed, the lesions in 31 were histologically confirmed by resection or biopsy. Seventeen patients were men, 14 were women. Ages ranged from 27 to 70, with a mean of 46 years. Primary (pre-treatment) lesions included low-grade astrocytoma (n = 3), malignant astrocytoma (n = 8), glioblastoma multiforme (GBM) (n = 12), anaplastic ganglioglioma (n = 1), and gliosis (n = 1). Four patients had astrocytomatas, the grade of which was not specified. The primary lesions in the remaining two patients were unknown. Average time interval between PET examination and surgery was 34 days. The longest interval was 220 days in a patient with a hypermetabolic lesion that surgical resection confirmed to be a GBM.

Results

The results of the 31 histologically confirmed cases are summarized in Tables 1 through 4. Twenty-two of the 31 patients had histologically confirmed tumor recurrence (Table 1). Of those, 19 (86%) had lesions hypermetabolic relative to normal contralateral white matter. Two patients (9%) with recurrent tumor had hypometabolic PET abnormalities (patients 5 and 77). Both were histologically proved to be high-grade lesions; one was a glioblastoma, the other a malignant astrocytoma. A single patient with a recurrent GBM had a PET abnormality that was interpreted as equivocal because the metabolic activity was only marginally increased relative to contralateral white matter (patient 50).

Eight of the 31 patients had histologically confirmed radiation necrosis (Table 2). The PET scan abnormality in seven (88%) of those eight cases was hypermetabolic compared with normal contralateral white matter. Only a single case (13%) of histologically verified radiation necrosis was hypometabolic. There was also one case of histologically proved gliosis in a patient with a treated anaplastic oligodendroglioma (patient 47) in which the PET lesion was hypometabolic compared with contralateral normal white matter.

In all, 26 of the 31 patients had hypermetabolic PET scan abnormalities (Table 3). Nineteen (73%) of those patients had recurrent tumor. However, seven (27%) of the 26 had radiation necrosis. Four patients had hypometabolic PET abnormalities, of whom two (50%) had recurrent tumor, one (25%) had radiation necrosis, and one had the aforementioned gliosis. Therefore, using contralateral white matter as the internal standard, the sensitivity for diagnosis of recurrent tumor was calculated to be 86% ± 14% (95% CI); the specificity was 22% ± 6%. Positive and negative predictive values were 73% ± 20% and 50% ± 49%, respectively.

To eliminate the possibility that sampling error or specimen inadequacy had biased the results, we considered the surgically resected cases (n = 19) separately from the lesions for which a biopsy had been done (Table 4). In 17 of those 19 cases, the abnormality on FDG PET scans was hypermetabolic relative to normal contralateral white matter. Twelve (71%) of those lesions were histologically proved to be recurrent tumor. Five (29%) of the hypermetabolic lesions were radiation necrosis. The single hypometabolic lesion in the surgically resected group was a malignant astrocytoma (patient 5). The patient with the recurrent GBM whose PET study was inter-
preted as equivocal (patient 50) also had the lesion resected.

Table 5 summarizes the results comparing metabolic activity of the FDG PET scan abnormality with normal contralateral gray matter. Of the 22 recurrent tumors, 16 (73%) were hypermetabolic and six (27%) were hypometabolic relative to normal gray matter. Metabolic activity in the eight cases of radiation necrosis was equally split between hypermetabolism and hypometabolism. In the sole case of gliosis, the tumor was also hypometabolic. On the basis of this comparison scheme, 16 (80%) of the 20 patients with hypermetabolic FDG PET scan abnormalities had recurrent tumor; the remaining 20% had radiation necrosis. Conversely, five (45%) of 11 patients with hypometabolic lesions had histologically proved radiation necrosis while 55% had recurrent tumor. Therefore, using gray matter as the standard of comparison, the calculated sensitivity of FDG PET for diagnosis of recurrent tumor was 73% ± 14% and the specificity was 56% ± 18%. Positive and negative predictive values were 80% ± 14% and 46% ± 14%, respectively.

Figures 1 and 2 illustrate the potential false-positive errors that can arise using FDG PET to distinguish recurrent tumor from radiation necrosis. In both cases, the lesions were hypermetabolic compared with normal contralateral white matter. Both findings were believed to be consistent with recurrent tumor; however, histologic evaluation confirmed both to be radiation necrosis.

Discussion

Differentiation between residual viable or recurrent tumor and radiation necrosis in patients with
treated primary or metastatic CNS tumors has historically been very difficult. Conventional contrast-enhanced CT and MR studies have proved very reliable for tumor detection. MR imaging is also effective for depicting the effects of radiation therapy on the brain (1, 5). However, neither CT nor MR imaging has been able to reliably distinguish recurrent tumor from radiation necrosis (1–5). This may be due at least in part to the limited number of ways the brain can respond to various insults: tumor and radiation necrosis can both incite vasogenic edema, disrupt the blood-brain barrier, and cause cavitation.

The use of functional FDG PET to distinguish recurrent tumor from radiation necrosis appeared to be promising on a theoretical basis. It had long been known that tumor cells have altered glucose metabolism owing to an increased reliance on anaerobic glycolysis (16). It was also known that irradiation of tumor cells caused glucose utilization to decrease (17). Because regions of radiation necrosis contain irradiated tumor cells, areas of coagulation necrosis, reactive gliosis, and active fibrosis, it was theorized that assessment of glucose utilization would be able to distinguish recurrent tumor from radiation necrosis. After Phelps and colleagues at UCLA found a reproducible method of measuring metabolic rates in vivo with FDG PET, it appeared there might finally be a reliable tool with which to diagnose recurrent tumor (18). Early work at the National Institutes of Health supported this optimism (6, 8). Patronas et al (6) used FDG to examine five patients with radiation-treated glial neoplasms. All three patients with hypermetabolic PET scan abnormalities were histologically shown to have recurrent tumor. The two patients with hypometabolic PET lesions had proved radiation necrosis. Recently, there has been a growing concern that FDG PET is more limited in its ability to distinguish recurrent tumor from radiation necrosis than initially believed (13, 14, 19). Our results support these recent studies.

Of the 31 histologically proved cases in this study, the FDG PET studies were correct in 21 (68%) and incorrect in nine (29%). In the remaining case (patient 50), the PET scan was interpreted as indeterminate because the metabolic activity was not sufficiently different from normal contralateral white matter to unequivocally suggest tumor recurrence. As previously noted, it was subsequently shown to be a GBM.

In the 21 cases in which the PET scan findings agreed with the histologic diagnosis, there were 19 hypermetabolic and two hypometabolic lesions compared with normal white matter. The hypermetabolic lesions were all histologically confirmed tumor recurrence. Both hypometabolic lesions were tumor free. One was histologically proved radiation necrosis; the other was simply gliosis in a patient who had had therapy for an anaplastic oligodendroglioma.

Nine patients in our series had lesions that were misdiagnosed on the basis of FDG PET findings. Two patients had false-negative PET results in which hypometabolic lesions were thought to represent radiation necrosis; both were histologically shown to be recurrent malignant tumor. According to early FDG PET studies, one would expect such high-grade lesions to be hypermetabolic (6–8). The remaining seven misdiagnoses were all false-positive hypermetabolic lesions that were subsequently proved to be radiation necrosis. The lesions in five of those patients were completely resected, so the results do not appear to be related to sampling error.

As expected, the majority (86%) of recurrent tumors in this study were hypermetabolic compared with normal contralateral white matter. Unfortunately, the majority of cases of radiation necrosis (88%) were also hypermetabolic. Overall, the chance that any given hypermetabolic PET lesion represented recurrent tumor was 73%, while the chance that any hypometabolic PET lesion did not represent recurrent tumor was 50%. False-negative PET findings in patients with recurrent high-grade tumors have been well described in the past and were not unexpected (13, 14). Furthermore, there have been reports of false-positive FDG PET findings, including one case in which the increased glucose utilization was related to subclinical seizure activity (9, 20).
However, the large number of false-positive interpretations, (seven of 26 or 27%) in this study, was surprising. To our knowledge, this is the largest series of false-positive FDG PET findings yet reported. The cause of the elevated metabolic activity in these regions of radiation necrosis is unclear. Because this was a retrospective study, it was not possible to examine the patients for subclinical seizure activity. However, it seems unlikely that the hypermetabolism in all seven patients was seizure induced.

As noted, the sensitivity of FDG PET for detection of recurrent tumor using contralateral normal white matter as the internal standard was 86%; the specificity was 22%. The positive and negative predictive values were 73% and 50%, respectively. Excluding the single indeterminate scan, nine (29%) of 31 patients would have been treated inappropriately had the PET scan been the sole determinant of additional therapy. Seven of those patients would have undergone unnecessary treatment for radiation necrosis while two patients with recurrent high-grade tumor would have had therapy withheld. When contralateral normal gray matter was used for comparison, the number of correctly diagnosed tumors decreased from 19 to 16; the number of correctly diagnosed nontumorous lesions increased from two to five (Table 5). Therefore, while comparison with contralateral gray matter improved the diagnosis of radiation necrosis, 14% more tumors were misdiagnosed. Furthermore, while the specificity of PET increased to 56%, the sensitivity decreased to 73%. Changes in the positive and negative predictive values were small. Finally, 10 (32%) of 31 patients would have been treated inappropriately had contralateral gray matter been used for comparison (versus 29% had contralateral white matter been used). Thus, even using gray matter as the comparison standard does not improve FDG PET results enough to make it a clinically reliable tool for differentiating recurrent tumor from radiation necrosis. These results, particularly the low
those patients would have undergone surgery. Second, if the decision to do a biopsy or clinical criteria and MR findings as well as PET results were the sole determinant of therapy, as was done in several prior studies, we would most likely have misdiagnosed several hypermetabolic lesions as recurrent tumor. Therefore, it can be argued that our sensitivity and specificity values are more accurate than those of prior studies, in which not all results were histologically verified.

A 1993 commentary by DiChiro and Fulham (24) suggested that FDG PET is a better long-range predictor of disease outcome in CNS tumor cases than histology (24). That may well be true of primary, untreated tumors. However, this study addressed a fundamentally different question in which different information was desired from the PET scan. There is no argument that the majority of treated high-grade glial tumors, particularly glioblastomas, will ultimately recur. However, when patients with treated tumors present with clinical or MR evidence of disease progression, the question we need to answer is: Do FDG PET scan results accurately reflect the underlying histology so that a reliable decision regarding additional therapy can be made? Our data suggest that FDG PET cannot reliably answer this question.

Conclusion

The results of this study suggest that the ability of FDG PET to differentiate recurrent tumor from radiation necrosis is limited. Both false-positive and false-negative PET results contributed to unacceptably low sensitivity, specificity, and negative predictive values. Although the specificity improved when contralateral gray matter was used as the comparison standard, the results remained inadequate as a basis for therapeutic decisions. Additionally, changing the comparison standard to gray matter did not significantly alter the percentage of patients who would have been treated inappropriately had PET scan results been the sole determinant of therapy. Thus, investigation into other PET imaging agents and alternative imaging methods is still needed.

References


Please see the Editorial on page 590 in this issue.