Monitoring Cancer Treatment with PET/CT: Does It Make a Difference?

Wolfgang A. Weber[1] and Robert Figlin[2]

[1]Department of Molecular and Medical Pharmacology, University of California, Los Angeles, California; and [2]Division of Medical Oncology and Therapeutics Research, City of Hope Comprehensive Cancer Center, Los Angeles, California

PET with the glucose analog ^{18}F-FDG is increasingly being used to monitor the effectiveness of therapy in patients with malignant lymphomas and a variety of solid tumors. The use of integrated PET/CT instead of stand-alone PET for treatment monitoring poses some methodologic challenges for the quantitative analysis of PET scans but also provides the opportunity to integrate morphologic information and functional information. This integration may allow the definition of new parameters for assessment of the tumor response and will also facilitate the use of PET in research studies as well as in clinical practice. This review addresses how CT-based attenuation correction may affect the quantitative analysis of ^{18}F-FDG PET scans and summarizes the results of recent studies with PET/CT for treatment monitoring for lung cancer and gastrointestinal stromal tumors. The review concludes with an outlook on how PET/CT could make a difference in drug development and clinical management for patients.

Key Words: PET; ^{18}F-FDG; CT; tumor response; treatment monitoring; patient outcome

J Nucl Med 2007; 48:36S–44S

Within the last 5 y, PET/CT (*1*) has almost replaced stand-alone PET for imaging of patients with cancer, and there is little doubt that in the near future, most oncologic PET will be PET/CT. PET/CT improves the anatomic localization of abnormalities identified by PET and reduces the number of false-positive studies by facilitating identification of the physiologic accumulation of ^{18}F-FDG in normal organs, such as skeletal muscles or the genitourinary tract. One source of false-positive findings with ^{18}F-FDG PET, metabolically active brown adipose tissue, had not even been identified until PET/CT became available. Several studies have now indicated that for a variety of cancers, PET/CT has significantly higher staging accuracy than PET or CT alone. Perhaps most importantly, the anatomic information provided by PET/CT has made it much easier to communicate ^{18}F-FDG PET findings to referring physicians and has markedly improved their confidence in the test results. In summary, tumor staging by PET/CT has many advantages and no apparent disadvantages compared with tumor staging by stand-alone PET.

The situation is different for the quantitative analysis of ^{18}F-FDG PET scans and treatment monitoring. For these purposes, PET/CT also offers several potential advantages, but there are also some technical challenges that may limit the accuracy of PET/CT for assessing treatment effects. The major concern is that CT-based attenuation correction may be inaccurate because of differences in the photon energies used for PET and CT, misregistration of the PET and CT datasets, and the administration of contrast agents.

In this review we first discuss potential methodologic challenges in the use of PET/CT for treatment monitoring. The second part of the review addresses how PET/CT may improve evaluation of the tumor response by combining anatomic information and functional information. Both parts attempt to address the question of whether data analysis and image interpretation are different for PET/CT and PET. In the third part of the review, we summarize the results of recent studies with PET/CT for treatment monitoring. Because the number of such studies is still very limited, it is currently not possible to draw firm conclusions about whether PET/CT makes a difference. Therefore, this review concentrates on general concepts for the use of PET/CT and not on specific results for individual tumor types. In other words, it addresses whether PET/CT can make a difference rather than answering the question of whether PET/CT does make a difference. Concepts for the use of PET/CT in drug development and clinical practice are discussed in the fourth part of this article. This part addresses whether monitoring tumor response by PET/CT imaging can make a difference in drug development or clinical practice. To avoid overlap with other, current reviews on treatment monitoring with PET (e.g., *2–5*), we do not aim to provide in this article an overview of the diagnostic performance of PET for assessing the tumor response in various diseases; instead, we specifically address how PET/CT may be integrated in drug development and disease management. Although many tracers may be used for PET/CT in the future, this review discusses only ^{18}F-FDG PET/CT,

Received Sep. 8, 2006; revision accepted Sep. 26, 2006.

For correspondence or reprints contact: Wolfgang A. Weber, MD, Department of Molecular and Medical Pharmacology, UCLA, 10833 Le Conte Ave., Los Angeles, CA 90095.

E-mail: wweber@mednet.ucla.edu

because there are currently only very limited published data on the use of PET/CT for treatment monitoring with tracers other than ^{18}F-FDG.

METHODOLOGIC CHALLENGES OF PET/CT

With PET/CT, attenuation maps generated from CT scans are generally used to correct PET emission scans for photon attenuation. CT-based attenuation correction provides several important advantages for clinical PET but also poses some methodologic challenges for the quantitative analysis of PET studies and monitoring of the effects of therapy.

With stand-alone PET, attenuation correction is generally achieved by transmission scans with positron-emitting radioactive sources (*6*). Because of the limited counting rate capability of PET detector systems, these scans need to be acquired over several minutes to obtain sufficient count statistics for the generation of accurate attenuation maps. The segmentation of transmission images has allowed a significant reduction in the duration of a transmission scan, but the duration of the transmission scan (about 2–3 min per bed position) still significantly contributes to the overall duration of a whole-body PET study. Furthermore, transmission data can be contaminated by the 511-keV photons emitted by the PET tracer, resulting in underestimation of the true photon attenuation in areas with high radioactivity concentrations, such as the bladder or tumors with very high levels of tracer uptake (*6*).

All of these limitations of attenuation correction with radioactive sources are eliminated by CT-based attenuation correction. A whole-body CT scan with a multidetector system lasts only a few seconds. The duration of a whole-body PET study can therefore almost be halved by CT-based attenuation correction. Furthermore, the much higher photon flux of CT reduces the noise of attenuation maps and eliminates underestimation of the photon attenuation caused by contamination of the transmission data with emission photons.

However, the much shorter duration of the CT scan can also result in significant misregistration of PET and CT images. Although PET images are averaged over several breathing cycles, CT scans are acquired during a single breathing cycle. Consequently, the positions of the lungs and liver are likely to differ between PET and CT images, resulting in incorrect attenuation correction of the PET emission data.

In addition, PET uses monoenergetic 511-keV photons, whereas the x-ray tube of a CT scanner emits a spectrum of photons with an effective energy of approximately 70 keV. Because of this difference in photon energies, the attenuation coefficients derived from CT images need to be scaled appropriately to be used for attenuation correction of PET images (*7*). This process is not straightforward, because the scaling factors are dependent on the tissue type; therefore, CT images need to be segmented before they can be used for attenuation correction of the PET data. Several studies have evaluated how misregistration of PET and CT images because of respiratory motion and scaling of the attenuation correction factors influences the qualitative and quantitative analyses of PET scans with ^{18}F-FDG. The results of these studies are briefly summarized in the following sections.

Misregistration of PET and CT Image Data

If a CT scan is acquired during maximum inspiration, then the position of the diaphragm on the CT images will be up to several centimeters lower than it is on the PET images. In this situation, if the CT scan is used to correct the PET emission data for photon attenuation, then the measured activity concentration in the upper regions of the liver will be underestimated, because photons are attenuated more significantly by liver tissue than by lung tissue (*8*). Acquiring CT scans during expiration significantly reduces the misregistration of PET and CT images, because during a normal respiratory cycle, the chest is in the expiratory position much longer than it is in the inspiratory position (*8*). However, the midexpiratory position of a normal breathing cycle may be difficult to reproduce voluntarily during CT acquisition. In 1 study with midexpiratory CT scans for attenuation correction, the misalignment of the diaphragm on PET and CT scans was more than 1 cm on 50% of the scans and more than 2 cm on 34% of the scans. For lung tumors ranging in size from 0.9 to 2.3 cm, this misalignment resulted in an underestimation of ^{18}F-FDG uptake by up to 50% (*9*).

Because of the difficulties in reproducing the average midexpiratory position during CT studies, many centers are currently using shallow free breathing for the acquisition of CT scans (*10*). The average respiratory movement of lung tumors during shallow breathing has been reported to be on the order of 2 mm in the mediolateral and anterior–posterior directions, whereas craniocaudal movement is, on average, 4 mm (*11*). However, the degree of respiratory movement varies significantly with the location of a tumor (*12*), with significantly more craniocaudal respiratory movement occurring in the lower lung fields (mean ± SD, 12 ± 6 mm) than in the upper lung fields (2 ± 2 mm). Larger lesions that are attached to the chest wall or vertebral bodies demonstrate only minimal movement. Nevertheless, for small pulmonary lesions, respiratory gating or respiratory averaging of CT scans can significantly improve the accuracy of quantitative measurements in ^{18}F-FDG PET/CT studies (*9,13*).

Scaling of CT-Based Attenuation Coefficients and Influence of Contrast Agents

The attenuation of photons by a given material depends on their energy and the density of the material. High-energy photons are attenuated less than low-energy photons, and high-density materials attenuate photons more efficiently than low-density materials. However, attenuation is also influenced by the effective atomic number (Z) of the material. At the energy levels of CT, photons primarily interact through photoelectric effects, whose likelihood is proportional

to Z^4. In contrast, the 511-keV annihilation photons used for PET are almost exclusively attenuated by Compton scattering, which shows little dependence on Z. This means that materials with a high Z attenuate CT photons much more efficiently than the 511-keV photons emitted during positron decay, whereas differences in photon attenuation are much smaller for materials with a low Z (*7*). For example, at 70 keV (the effective mean energy of photons emitted by a 140-kVp x-ray source), the mass attenuation coefficient of water is 1.8 times higher than it is at 511 keV. For calcium, however, the attenuation coefficient is 4 times higher at 70 keV than at 511 keV. For iodine, the ratio of the attenuation coefficients at 70 and 511 keV is more than 36 (*7,14*). Therefore, it is not feasible to scale CT images with a single factor to generate a correct attenuation map for 511-keV photons. It is necessary to segment a reconstructed CT scan into different tissue types (with different Z values) and then use the appropriate scaling factor for each individual tissue type. Errors in the segmentation process cause incorrect attenuation correction factors and result in incorrect quantitative data in the attenuation-corrected PET emission scan.

Studies have shown that it is sufficient to segment CT images into 3 tissue types (air, water, and bone) to generate an accurate attenuation map for PET from CT images (*7,15*). However, in the presence of CT contrast agents, this approach results in incorrect attenuation correction factors, because the attenuation of the 511-keV photons by the contrast agents is overestimated as a result of their high Z values and the resulting high level of attenuation of the x-ray photons (*16*). Therefore, the true activity concentration in contrast agent–filled blood vessels is overestimated (*16*). In contrast-enhancing tumors, a similar effect has been observed, but studies have suggested that, on average, the increase in the measured tumor ^{18}F-FDG uptake is small (*14,17–19*). Nevertheless, in individual lesions, larger variations have been reported (*14,19*). For accurate measurements of tumor ^{18}F-FDG uptake, it may therefore be preferable to acquire a low-dose non–contrast-enhanced CT scan for attenuation correction before the PET emission scan is acquired and to obtain a separate, contrast-enhanced (diagnostic) CT scan after the PET emission scan is obtained.

Combining Anatomic Information and Functional Information with PET/CT

Quantitative parameters derived from PET studies are generally based on activity concentrations. For example, standardized uptake values (SUVs) describe the ratio between the activity concentration in the tumor and the (hypothetical) activity concentration in the whole body, if the radioactivity were homogeneously distributed throughout the patient. Similarly, metabolic rates describe the amount of glucose metabolized per gram of tissue per unit of time. Therefore, tumor growth or shrinkage does not necessarily affect SUVs or metabolic rates, because a tumor could grow or shrink without a change in the metabolic rate per gram of tissue. As a consequence, SUVs or metabolic rates only incompletely capture the tumor response to therapy, because they do not reflect total tumor metabolic activity but instead reflect metabolic activity per gram of tissue.

To overcome this limitation, Larson et al. proposed several years ago (*20*) that the metabolic rate or SUV of a lesion be multiplied by its volume to obtain "total lesion glycolysis." In untreated tumors with high levels of glucose metabolic activity, it is frequently straightforward to estimate tumor volumes in PET by including voxels with an activity concentration above a certain threshold in the tumor volume. During or after treatment, however, it can be challenging to determine tumor volumes in PET, because contrast between the tumor and surrounding normal tissues is frequently low and no clear tumor borders can be identified. With integrated PET/CT, it is now possible to measure the tumor volume in CT and multiply this volume by the SUV measured in PET to obtain total lesion glycolysis.

Multiplying tumor volume and ^{18}F-FDG uptake is only 1 of several approaches to combining morphologic information and anatomic information for a better assessment of the tumor response. For example, it has been shown for esophageal cancer that changes in the tumor volume 2 wk after the start of chemotherapy are significantly correlated with the histopathologic tumor response (*21,22*). Therefore, it will be important to determine in future studies whether early metabolic and volumetric changes provide independent prognostic information and whether the accuracy of response prediction by ^{18}F-FDG PET can be improved by defining response criteria that are based on metabolic and volumetric changes. Integrated PET/CT greatly facilitates such studies (*23*), because PET and CT scans are acquired during the same imaging session (Fig. 1).

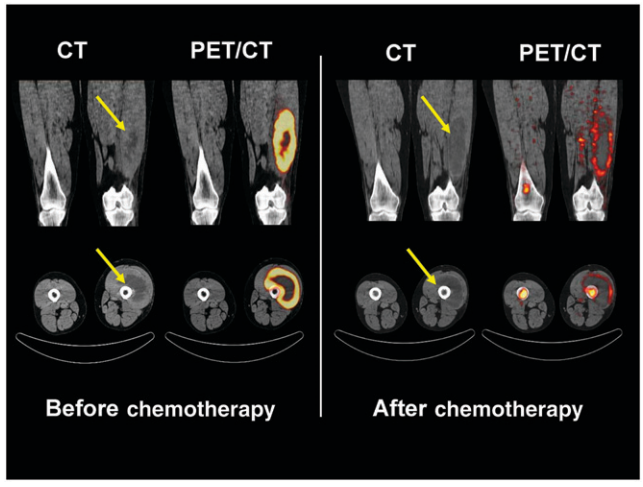

FIGURE 1. ^{18}F-FDG PET/CT studies in patient with soft-tissue sarcoma of right thigh. Patient was treated with presurgical chemotherapy, and histopathologic analysis revealed 95% treatment-induced necrosis. This finding was reflected by marked decrease in tumor ^{18}F-FDG uptake. Quantitatively, tumor SUV decreased from 10.0 to 1.0. In contrast, there was no major change in tumor size on CT (arrows).

In addition to measurements of changes in tumor size, CT information may be used to improve partial-volume correction of quantitative parameters derived from PET images. Currently, the most commonly used approach for partial-volume correction is to assume that the tumor is approximately spheric, demonstrates homogeneous tracer uptake, and is located in a homogeneous background. Under these assumptions, the "true" radioactivity concentration in the tumor can be calculated by dividing the background-corrected activity concentration in the tumor by the recovery coefficient of a sphere of the same size (*24*). These simple assumptions may be appropriate in some clinical situations, such as a solitary pulmonary nodule surrounded on all sides by lung tissue. Frequently, however, tumors have an irregular shape and background activity is heterogeneous. In a patient with a pulmonary mass located close to the mediastinum, defining background activity can be quite arbitrary, because one can choose either the mediastinal blood pool or the lung to represent the background. Under these circumstances, the validity of the simple partial-volume correction algorithm described earlier can be questioned. With the use of coregistered CT images, more sophisticated partial-volume correction methods (*25*) that may improve quantitative measurements of tracer concentrations in tumors can be developed.

CT images may also be helpful for placing regions of interest (ROIs) to analyze PET studies more accurately and reproducibly. Currently, various approaches are being used to define ROIs on ^{18}F-FDG PET scans. These include manually contouring the outer borders of the tumor in 1 slice or several slices, using semiautomated algorithms based on threshold values, or placing a fixed ROI in the area of the tumor with the highest level of ^{18}F-FDG uptake. Because all of these approaches are, to some extent, user dependent, some researchers have advocated using the pixel with the maximum SUV for data analysis. However, the maximum SUV is the most sensitive to statistical noise as well as image reconstruction and postprocessing parameters. Using the anatomic information from CT images to define tumor extent and copying the ROIs to PET images may allow a more reproducible definition of ROIs.

All of these approaches to integrating functional information and morphologic information provide exciting opportunities for improving treatment monitoring in patients with cancer. However, it is probably at least equally important that PET/CT greatly facilitates the integration of treatment monitoring with PET in clinical practice and medical research. With stand-alone PET, PET and CT scans are performed separately, frequently at different institutions. This scenario makes timely scheduling of follow-up studies complex, increases the time burden for patients, and may result in conflicting image interpretations. Integrated PET/CT helps to avoid these problems, because PET and CT scans for assessment of the tumor response can be obtained in 1 imaging session and the results of PET and CT can be jointly reported. These seemingly trivial but important practical factors are likely to considerably increase the acceptance of PET as a tool for monitoring the tumor response to therapy.

Examples of Treatment Monitoring with PET/CT

Pöttgen et al. (*26*) recently evaluated PET/CT for monitoring presurgical therapy in patients with locally advanced non–small cell lung cancer (NSCLC). PET/CT scans were acquired during shallow breathing, and no respiratory gating was used. Fifty consecutive patients with stage IIIA/IIIB NSCLC treated by presurgical chemoradiotherapy were retrospectively analyzed. Patients underwent 3 cycles of platinum-based induction chemotherapy followed by chemoradiotherapy (total dose, 44–45 Gy). Pretreatment PET/CT scans were obtained about 3 d before the initiation of therapy and again after induction chemotherapy as well as after the completion of chemoradiotherapy. Tumors in 37 patients were considered to be resectable after chemoradiotherapy, and these patients underwent thoracotomy. Tumors with more than 90% necrosis were classified as responding. The ^{18}F-FDG uptake of the primary tumor and metastatic lymph nodes was quantified on the basis of the maximum SUV. ^{18}F-FDG uptake was corrected for partial-volume effects by measuring lesion size on CT and dividing the background corrected maximum SUV by the recovery coefficient of a sphere with the same diameter.

Relative changes in tumor ^{18}F-FDG uptake from the baseline scan to the first follow-up scan (after the completion of chemotherapy) were highly significantly correlated with the histopathologic response (area under the receiver operating characteristic curve, 0.88; $P = 0.005$). Changes in tumor ^{18}F-FDG uptake from the baseline scan to the second follow-up scan (after chemoradiotherapy) had a similar predictive value (area under the receiver operating characteristic curve, 0.86; $P = 0.008$). At this time, a relative decrease in tumor ^{18}F-FDG uptake of between 45% and 62% predicted the histopathologic response with sensitivities ranging from 94% to 70% and specificities ranging from 86% to 71%.

The study of Pottgen et al. (*26*) indicates that in patients with locally advanced NSCLC, quantitative analysis of routine whole-body ^{18}F-FDG PET/CT studies is feasible and allows an accurate prediction of the tumor response. The reported diagnostic accuracy for assessment of the histopathologic response was comparable or superior to that in a previous study evaluating stand-alone PET for patients with locally advanced NSCLC and undergoing presurgical therapy (*27*). The study of Pottgen et al. (*26*) also illustrates the use of the CT information from a PET/CT study to correct the measured tumor ^{18}F-FDG uptake for partial-volume effects.

Two studies (*28,29*) evaluated the use of ^{18}F-FDG PET/CT for monitoring the treatment of gastrointestinal stromal tumors (GIST) with the tyrosine kinase inhibitor imatinib (Gleevec; Novartis Pharma). Both confirmed the finding of previous studies with ^{18}F-FDG PET that tumor ^{18}F-FDG uptake rapidly decreases during treatment with imatinib.

Furthermore, they demonstrated how assessment of the tumor response can be improved by in-line PET/CT for patients with multiple metastatic lesions. The ^{18}F-FDG uptake of untreated GIST is variable, with some tumors demonstrating only mildly increased metabolic activity. In a study by Antoch et al. (*29*) of 20 patients with GIST, PET detected 135 lesions, whereas CT revealed 249. Lesions missed by PET included intraabdominal deposits that were incorrectly classified as ^{18}F-FDG uptake in normal bowel as well as pulmonary metastases. Integrated PET/CT detected 282 lesions ($P < 0.001$). In a study by Goerres et al. (*28*) of 15 patients with GIST, 66 lesions were identified by PET, whereas 96 lesions were found by CT. Because not all metastatic lesions were found by stand-alone PET at baseline and new PET-negative lesions developed in some patients during treatment, the response assessment by in-line PET/CT was overall more accurate than that by PET alone. Conversely, the response assessment by CT was limited by the fact that tumor shrinkage in response to imatinib therapy was slow. Therefore, Antoch et al. (*29*) used the following algorithm to combine tumor response assessments by PET and CT. The tumor response was assessed by PET according to the criteria of the European Organization for Research and Treatment of Cancer (EORTC) (*30*), whereas the response assessment by CT was based on response evaluation criteria for solid tumors (RECIST). If either CT or PET data indicated progressive disease, then the response was classified as progressive disease. If 1 reading was no change and the other was a complete response, then the consensus was designated a partial response. If either CT or PET data suggested no change but the other imaging modality indicated a partial response, then the final decision was based on the density of the lesion on CT, as measured by Hounsfield units (HU). Decreasing HU indicated a response, whereas a lack of change in HU was interpreted as no change. With these criteria, the tumor response could be assessed by in-line PET/CT with an overall accuracy of 95% after 1 mo of gefitinib therapy. The accuracies of PET and CT alone were only 85% and 44%, respectively. In summary, the studies of Antoch et al. and Goerres et al. (*28,29*) demonstrated how response criteria integrating morphologic information and functional information can be defined. Furthermore, they showed that assessment of the tumor response by CT is not limited only to measurements of tumor size, because changes in HU can provide additional information.

CONCEPTS FOR TREATMENT MONITORING WITH PET/CT IN DRUG DEVELOPMENT AND CLINICAL PRACTICE

Need for New Tools to Monitor Tumor Response to Therapy

Assessment of the tumor response to therapy plays a central role in drug development as well as in clinical management for patients. Currently, the response is mainly evaluated by measuring tumor size with CT and classifying tumor shrinkage according to standard criteria, such as those of the World Health Organization (WHO) or RECIST (*31,32*). The original WHO response criteria were based on bidimensional measurements of the tumor and defined response as a decrease of at least 50% in the sum of the product of the longest perpendicular diameters of measured lesions. The rationale for using a 50% threshold for the definition of response was based on data evaluating the reproducibility of measurements of tumor size by palpation and on planar chest radiographs (*31,33*). The more recent RECIST introduced by the National Cancer Institute and the EORTC criteria standardized imaging techniques for anatomic response assessment by specifying minimum requirements for acquisition parameters in CT and providing size thresholds for measurable lesions. In addition, in RECIST, the longest bidirectional diameters were replaced with the longest unidimensional diameter as the representation of a measured lesion (*32*). RECIST defines response as a 30% decrease in the largest diameter of a tumor. For a spheric lesion, this value is equivalent to a 50% decrease in the product of 2 diameters.

Metaanalyses combining the results of several large phase II and phase III studies have shown that the tumor response based on WHO criteria or RECIST is correlated with patient survival for some tumor types (*34*). However, there is considerable variability between individual studies, and the same response rate can be associated with completely different survival rates in different studies (*35*). For some tumor types, metaanalyses found no or only a very weak correlation with patient survival (*36,37*). Given the history of response criteria outlined earlier, these observations are probably not unexpected. Current response criteria were designed to ensure a standardized classification of tumor shrinkage in response to therapy. They were not developed on the basis of clinical trials correlating tumor shrinkage with patient outcome. Although tumor shrinkage may generally be expected to be associated with a better outcome of therapy, it is also clear that there are disease- and treatment-specific differences. For example, a certain degree of tumor shrinkage in a patient with NSCLC may indicate a relatively good prognosis in comparison with that for other patients with NSCLC. However, the same degree of tumor shrinkage may be associated with a relatively poor prognosis in a patient with Hodgkin's disease, because Hodgkin's disease is much more sensitive to chemotherapy and, consequently, tumor shrinkage is, on average, much more pronounced in Hodgkin's disease than in NSCLC.

Furthermore, morphologic alterations may not be adequate for evaluating the response to newer cytostatic agents, with which anatomic changes may be absent or slow to be manifested. This situation is perhaps best illustrated by 2 large phase III trials evaluating epidermal growth factor receptor (EGFR) kinase inhibitors in patients with advanced NSCLC (*38,39*). In a study of 731 patients (*39*), the overall response rate for patients receiving the EGFR kinase inhibitor erlotinib (Tarceva; Genentech/OSI Pharmaceuticals)

was only 8.9%. Nevertheless, erlotinib improved the median overall survival rate by more than 50% compared with that obtained with a placebo ($P < 0.001$). A similar study evaluating the EGFR kinase inhibitor gefitinib (Iressa; AstraZeneca) included 1,692 patients with advanced NSCLC (*38*). In that study, the overall response rate was 8.2%. Thus, if response rates were good predictors of patient survival, then gefitinib treatment should result in a survival benefit similar to that obtained with erlotinib, because the response rates for erlotinib and gefitinib (8.9% and 8.2%, respectively) were almost identical. However, this was not the case. In contrast to erlotinib, gefitinib did not improve the survival rate compared with that obtained with a placebo (hazard ratio, 0.89; $P = 0.08$). The results of these 2 studies demonstrated that response rates for targeted drugs may be low despite significant clinical benefit and that response rates may be poor predictors of patient survival.

Currently, the response rate in phase II studies is the major criterion for whether or not to further pursue a drug candidate in large phase III trials. At least for EGFR kinase inhibitors, it is now unclear what a certain response rate means for the efficacy of a drug. For example, if a new EGFR kinase inhibitor were to show a response rate of 9% in a phase II study, one could interpret this as a positive result because erlotinib has been shown to significantly improve patient survival at this response rate. However, one could also argue that this response rate does not justify further testing of the new drug because gefitinib has been shown to be ineffective at this response rate.

Promise of ^{18}F-FDG PET/CT for Monitoring Tumor Response in Clinical Trials

Several studies have suggested that quantitative changes in tumor ^{18}F-FDG uptake 2–3 wk after the start of therapy have been shown to correlate well with subsequent tumor shrinkage and patient survival. Thus, ^{18}F-FDG PET has the potential to improve disease management by signaling the need for early therapeutic changes in nonresponders, thereby avoiding the side effects and costs of ineffective treatment. Furthermore, as an early indicator of clinical benefit, ^{18}F-FDG PET may also facilitate oncologic drug development by shortening phase II trials and detecting clinical benefit earlier in phase III investigations.

^{18}F-FDG PET is also attractive for monitoring treatment with certain protein kinase inhibitors, because many signaling pathways targeted by protein kinase inhibitors also have a well-established role in regulating tumor glucose metabolism. For example, the protein kinase Akt is a central regulator of cellular apoptosis (*40,41*) but is also involved in the regulation of glucose use (*42*). Recent experimental data suggest that the activation of Akt may be a key factor for the markedly induced glucose use of cancer cells (*43,44*).

As mentioned earlier, ^{18}F-FDG PET has already been used in clinical studies to monitor the response of GIST to treatment with imatinib (*29,45–47*). A marked reduction in tumor metabolic activity was noted as early as 24 h after the first dose of imatinib (*45,48*). Moreover, extensive anatomic abnormalities observed by CT persisted at a time when metabolic alterations had already resolved. This rapid change in ^{18}F-FDG uptake appears to be mediated by the translocation of glucose transporters from the plasma membrane to the cytosol and precedes cell death (*49*). These data suggest that ^{18}F-FDG PET may become a valuable tool for monitoring treatment with imatinib and potentially other protein kinase inhibitors. For example, experimental studies have shown that treatment with EGFR kinase inhibitors results in a rapid inhibition of glucose transport that precedes cell death or growth inhibition in sensitive tumors (*50,51*).

Changes in Disease Management Based on ^{18}F-FDG PET/CT

Currently, treatment monitoring with ^{18}F-FDG PET probably has the highest impact on disease management for malignant lymphomas. Several studies have indicated that patients with metabolically active residual masses after the completion of chemotherapy have a poor prognosis compared with patients with nonmetabolic residual masses (*4*). On the basis of these data, it has been proposed that ^{18}F-FDG PET be integrated into the International Workshop Criteria for assessment of the response of aggressive non-Hodgkin's lymphoma (*52*). At many centers, PET/CT is now frequently used to assess tumor viability in patients with residual masses after the completion of chemotherapy to guide additional therapy, such as radiotherapy. Furthermore, ^{18}F-FDG PET/CT is often performed after 2 cycles of chemotherapy to assess the tumor response. The goal is to tailor the intensity and type of treatment to the individual patient's prognosis to achieve cure with the least possible toxicity by minimizing the treatment for patients with a good prognosis and intensifying the treatment for patients with a poor prognosis (*4*).

For solid tumors, ^{18}F-FDG PET has shown similar encouraging data for the early differentiation of patients with a favorable prognosis from those with an unfavorable prognosis (*27*). However, for most solid tumors, ^{18}F-FDG uptake does not decrease as rapidly as it does for malignant lymphomas. As a consequence, quantitative measurements of tumor ^{18}F-FDG uptake are helpful for differentiating between responding and nonresponding tumors, whereas visual analysis is frequently sufficient for the evaluation of ^{18}F-FDG PET scans for patients with lymphomas, because the metabolic activity of responding lymphomas decreases to background levels within the first few chemotherapy cycles. Despite encouraging data, ^{18}F-FDG PET/CT currently has less impact on the management of solid tumors than it does on that of malignant lymphomas, mainly because the effectiveness of second-line therapies is currently limited. In the absence of therapeutic options, improved assessment of the tumor response will not substantially change disease management. However, this situation is likely to change in the

future, when more second-line therapies, such as protein kinase inhibitors, become available.

Already, an early assessment of the tumor response has the potential to guide the presurgical chemotherapy or chemoradiotherapy of solid tumors. Presurgical or neoadjuvant therapy is frequently used to decrease tumor size and improve resectability in patients with locally advanced disease. However, for several tumor types, such as esophageal or lung cancer, the impact of presurgical therapy on patient survival is small. Importantly, many studies have demonstrated that not all patients benefit a little from presurgical therapy but that a subgroup of patients responds very well and has excellent survival after neoadjuvant therapy and surgical resection. The majority of patients does not respond to therapy, and their survival is not better than that after surgical resection alone. Therefore, early identification of nonresponding tumors would be highly beneficial to avoid unnecessary toxicity and costs (*53*).

The MUNICON trial recently showed that quantitative measurements of tumor ^{18}F-FDG uptake can be used to individualize neoadjuvant therapy (*54*). This trial included patients with locally advanced adenocarcinomas of the esophagogastric junction and scheduled to undergo presurgical chemotherapy followed by surgical resection. ^{18}F-FDG PET was performed before therapy and after a short, 2-wk course of cisplatinum-based chemotherapy. If tumor ^{18}F-FDG uptake decreased by more than 35% at the time of the second PET scan, the patients underwent the full, 3-mo course of chemotherapy. Otherwise, the patients underwent immediate tumor resection. An interim analysis of this study confirmed that ^{18}F-FDG PET allows the selection of patients with a high probability of a histopathologic response (*54*).

Future Validation of ^{18}F-FDG PET/CT for Monitoring Tumor Response

Currently, there is no generally accepted definition for a metabolic response in ^{18}F-FDG PET. The EORTC published preliminary criteria for assessment of the tumor response in 1999 (*30*). However, at that time, only a limited number of data on the use of ^{18}F-FDG PET for treatment monitoring were available. Since then, a significant number of studies on treatment monitoring with ^{18}F-FDG PET have been published, and there is now a need to standardize the criteria used for monitoring anticancer therapy with ^{18}F-FDG PET.

On the basis of the current data regarding the test–retest reproducibility of ^{18}F-FDG PET, a 20% decrease in tumor ^{18}F-FDG uptake appears to be a reasonable working definition for a metabolic response. However, this definition of a metabolic response should be reevaluated in a multicenter setting, because the current data on test–retest reproducibility are based on 2 small single-center studies performed several years ago (*55,56*). Furthermore, it should be noted that a high test–retest reproducibility of ^{18}F-FDG PET can only be achieved if the scans are acquired and analyzed according to a strict protocol. Specifically, patients must be examined in the fasting state, baseline and follow-up studies must be acquired at the same time after ^{18}F-FDG injection, and blood glucose levels must be stable.

The National Cancer Institute recently published recommendations for acquiring ^{18}F-FDG PET studies in clinical trials; these recommendations describe in detail how patient preparation and the timing of data acquisition affect the quantitative assessment of tumor glucose metabolism by ^{18}F-FDG PET (*57*). These recommendations will make quantitative measurements of tumor ^{18}F-FDG uptake more consistent across different sites. However, further standardization of image acquisition and reconstruction parameters still appears to be necessary. This is not a trivial task, because PET detector technology and image reconstruction algorithms are constantly evolving. Therefore, recommendations for standardized parameters will quickly become outdated. Furthermore, vendor-specific differences in detector technology and image reconstruction algorithms may make it impossible to define parameters that are applicable across different scanner types. Consequently, it will be challenging to standardize the measurement process across different generations of PET scanners and across scanners from different manufacturers. It may be more feasible to standardize the results of the measurements. Instead of mandating specific acquisition and reconstruction parameters, the standard would request that measurements of activity concentrations in a phantom be within certain limits. Such an approach for standardization would be much more flexible, because image acquisition and reconstruction parameters could be scanner specific, as long as the results of the phantom measurements are comparable to those obtained with other scanner types.

A metabolic response defined on the basis of the test–retest reproducibility of ^{18}F-FDG PET will only describe the minimum effect of therapy that can be reliably determined in an individual patient. Thus, the definition is independent of specific tumor types or specific forms of therapy and may therefore be most appropriate for clinical trials evaluating new anticancer drugs. For established forms of therapy, it may be feasible to optimize response criteria. For example, in patients with high-grade malignant lymphomas, a mean decrease in ^{18}F-FDG uptake of more than 45% has been observed within 24 h after the administration of the first dose of chemotherapy (*58*). In patients with solid tumors treated by presurgical chemotherapy, a change in ^{18}F-FDG uptake of 30%–35% within the first few weeks of chemotherapy has been found to provide the highest accuracy for the prediction of histopathologically complete or subtotal tumor regression (*27*). Further validation of criteria for a metabolic response will require multicenter trials demonstrating that quantitative measurements of tumor ^{18}F-FDG uptake are robust and can be obtained with similar accuracies at multiple institutions.

One goal of such trials will be to determine whether changes in tumor ^{18}F-FDG uptake during therapy may be used as a surrogate end point in clinical trials. A surrogate end point must correlate with the true clinical outcome and fully

capture the net effect of treatment on the clinical outcome (*59*). Thus, studies demonstrating a strong correlation between changes in ^{18}F-FDG uptake and patient survival are necessary but not sufficient to establish a metabolic response in PET as a surrogate end point (*60*). A new form of therapy may result in a clinical benefit without causing a metabolic response early in the course of therapy, whereas others may induce a metabolic response not associated with a favorable clinical outcome. A frequently cited example for the difficulties in using surrogate end points in clinical trials is the use of bone mineral density measurements for assessment of the efficacy of drugs for the treatment of osteoporosis. It has been established that a low bone mineral density is associated with an increased risk of fractures. Some forms of treatment, such as bisphosphonates, increase bone mineral density and decrease the risk of fractures (*61*). In contrast, treatment with fluoride increases bone mineral density but does not decrease the risk of fractures (*62*). Thus, although a loss of bone mineral density is correlated with osteoporosis, bone mineral density is not a useful surrogate end point for monitoring treatment with fluoride.

Therefore, a series of randomized trials will be necessary to determine whether a metabolic response can be used as a surrogate end point in clinical trials. The general concept would be to include ^{18}F-FDG PET in several randomized phase III trials comparing the efficacies of a class of chemotherapeutic drugs for a certain disease. A consistent association of higher metabolic response rates with improved patient survival in these trials would provide evidence that a metabolic response may be used as a surrogate end point in this disease and for this class of drugs.

The second goal will be to develop strategies for the individualization of tumor therapy by ^{18}F-FDG PET. Figure 2 shows a potential design of trials addressing this question. Patients would be randomized to receive either standard or "PET-controlled" therapy. Patients undergoing PET-controlled therapy would receive standard therapy for a brief period of time (for example, 1 treatment cycle). Then the treatment response would be assessed by ^{18}F-FDG PET. Patients classified as metabolic responders would continue to receive the standard treatment, whereas an alternative treatment would be used for metabolic nonresponders. Examples of alternative treatments are immediate surgical resection instead of continued neoadjuvant therapy, the use of different, second-line, chemotherapy regimens, or the use of targeted drugs, such as protein kinase inhibitors. The end point of these studies would be costs, treatment-induced morbidity, and overall survival (Fig. 2).

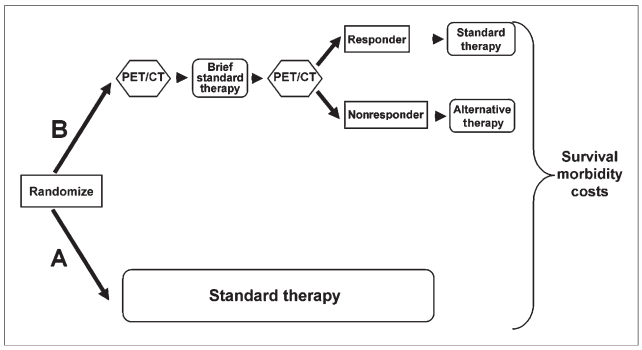

FIGURE 2. Proposed design of randomized trials evaluating impact of PET on disease management. Patients were randomized to either standard therapy (A) or PET-controlled therapy (B). For group B, treatment was changed on basis of tumor response after brief period of therapy. Metabolic responders (as determined by PET) continued to receive standard therapy, whereas treatment was changed for metabolic nonresponders. Effect of this change in therapy was evaluated by comparing survival, morbidity, and treatment-related costs in arms A and B.

CONCLUSION

PET/CT has the potential to make a difference in monitoring the tumor response to therapy by integrating anatomic and functional measures of treatment effects. Perhaps even more importantly, PET/CT is likely to increase the acceptance of PET as a tool for assessing the tumor response in medical research and clinical practice, because anatomic and functional measurements can be obtained in just 1 imaging session and jointly reported. This property will facilitate randomized multicenter studies validating the impact of treatment monitoring with PET/CT on disease management for larger groups of patients.

ACKNOWLEDGMENTS

This study was supported by UCLA ICMIC (NIH P50 CA86306), UCLA Lung SPORE (NIH P50 CA0038), and the UCLA Institute of Molecular Medicine (DE-FC03-87E60615).

REFERENCES

1. Beyer T, Townsend DW, Brun T, et al. A combined PET/CT scanner for clinical oncology. *J Nucl Med.* 2000;41:1369–1379.
2. Weber WA. Positron emission tomography as an imaging biomarker. *J Clin Oncol.* 2006;24:3282–3292.
3. Juweid ME, Cheson BD. Positron-emission tomography and assessment of cancer therapy. *N Engl J Med.* 2006;354:496–507.
4. Mikhaeel NG. Use of FDG-PET to monitor response to chemotherapy and radiotherapy in patients with lymphomas. *Eur J Nucl Med Mol Imaging.* 2006; 33(suppl 13):22–26.
5. Westerterp M, van Westreenen HL, Reitsma JB, et al. Esophageal cancer: CT, endoscopic US, and FDG PET for assessment of response to neoadjuvant therapy—systematic review. *Radiology.* 2005;236:841–851.
6. Bailey DL. Transmission scanning in emission tomography. *Eur J Nucl Med.* 1998;25:774–787.
7. Kinahan PE, Townsend DW, Beyer T, Sashin D. Attenuation correction for a combined 3D PET/CT scanner. *Med Phys.* 1998;25:2046–2053.
8. Goerres GW, Burger C, Kamel E, et al. Respiration-induced attenuation artifact at PET/CT: technical considerations. *Radiology.* 2003;226:906–910.
9. Pan T, Mawlawi O, Nehmeh SA, et al. Attenuation correction of PET images with respiration-averaged CT images in PET/CT. *J Nucl Med.* 2005;46:1481–1487.
10. Beyer T, Antoch G, Muller S, et al. Acquisition protocol considerations for combined PET/CT imaging. *J Nucl Med.* 2004;45(suppl 1):25S–35S.
11. Ekberg L, Holmberg O, Wittgren L, Bjelkengren G, Landberg T. What margins should be added to the clinical target volume in radiotherapy treatment planning for lung cancer? *Radiother Oncol.* 1998;48:71–77.

12. Seppenwoolde Y, Shirato H, Kitamura K, et al. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol Biol Phys.* 2002;53:822–834.
13. Nehmeh SA, Erdi YE, Pan T, et al. Four-dimensional (4D) PET/CT imaging of the thorax. *Med Phys.* 2004;31:3179–3186.
14. Berthelsen AK, Holm S, Loft A, Klausen TL, Andersen F, Hojgaard L. PET/CT with intravenous contrast can be used for PET attenuation correction in cancer patients. *Eur J Nucl Med Mol Imaging.* 2005;32:1167–1175.
15. Burger C, Goerres G, Schoenes S, Buck A, Lonn AH, Von Schulthess GK. PET attenuation coefficients from CT images: experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients. *Eur J Nucl Med Mol Imaging.* 2002;29:922–927.
16. Nakamoto Y, Chin BB, Kraitchman DL, Lawler LP, Marshall LT, Wahl RL. Effects of nonionic intravenous contrast agents at PET/CT imaging: phantom and canine studies. *Radiology.* 2003;227:817–824.
17. Yau YY, Chan WS, Tam YM, et al. Application of intravenous contrast in PET/CT: does it really introduce significant attenuation correction error? *J Nucl Med.* 2005;46:283–291.
18. Visvikis D, Costa DC, Croasdale I, et al. CT-based attenuation correction in the calculation of semi-quantitative indices of [^{18}F]FDG uptake in PET. *Eur J Nucl Med Mol Imaging.* 2003;30:344–353.
19. Mawlawi O, Erasmus JJ, Munden RF, et al. Quantifying the effect of IV contrast media on integrated PET/CT: clinical evaluation. *AJR.* 2006;186:308–319.
20. Larson SM, Erdi Y, Akhurst T, et al. Tumor treatment response based on visual and quantitative changes in global tumor glycolysis using PET-FDG imaging: the visual response score and the change in total lesion glycolysis. *Clin Positron Imaging.* 1999;2:159–171.
21. Beer AJ, Wieder HA, Lordick F, et al. Adenocarcinomas of esophagogastric junction: multi-detector row CT to evaluate early response to neoadjuvant chemotherapy. *Radiology.* 2006;239:472–480.
22. Wieder HA, Beer AJ, Lordick F, et al. Comparison of changes in tumor metabolic activity and tumor size during chemotherapy of adenocarcinomas of the esophagogastric junction. *J Nucl Med.* 2005;46:2029–2034.
23. Evilevitch V, Eilber F, Allen-Auerbach M, Phelps M, Czernin J, Weber W. Monitoring neoadjuvant chemotherapy of bone and soft-tissue sarcomas with FDG PET/CT [abstract]. *J Nucl Med.* 2006;47(suppl):103P.
24. Kessler RM, Ellis JR Jr, Eden M. Analysis of emission tomographic scan data: limitations imposed by resolution and background. *J Comput Assist Tomogr.* 1984;8:514–522.
25. Boussion N, Hatt M, Lamare F, et al. A multiresolution image based approach for correction of partial volume effects in emission tomography. *Phys Med Biol.* 2006;51:1857–1876.
26. Pottgen C, Levegrun S, Theegarten D, et al. Value of ^{18}F-fluoro-2-deoxy-D-glucose-positron emission tomography/computed tomography in non-small-cell lung cancer for prediction of pathologic response and times to relapse after neoadjuvant chemoradiotherapy. *Clin Cancer Res.* 2006;12:97–106.
27. Weber WA, Wieder H. Monitoring chemotherapy and radiotherapy of solid tumors. *Eur J Nucl Med Mol Imaging.* 2006;33(suppl 13):27–37.
28. Goerres GW, Stupp R, Barghouth G, et al. The value of PET, CT and in-line PET/CT in patients with gastrointestinal stromal tumours: long-term outcome of treatment with imatinib mesylate. *Eur J Nucl Med Mol Imaging.* 2005;32:153–162.
29. Antoch G, Kanja J, Bauer S, et al. Comparison of PET, CT, and dual-modality PET/CT imaging for monitoring of imatinib (STI571) therapy in patients with gastrointestinal stromal tumors. *J Nucl Med.* 2004;45:357–365.
30. Young H, Baum R, Cremerius U, et al. Measurement of clinical and subclinical tumour response using F-18-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. *Eur J Cancer.* 1999;35:1773–1782.
31. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer.* 1981;47:207–214.
32. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst.* 2000;92:205–216.
33. Moertel CG, Hanley JA. The effect of measuring error on the results of therapeutic trials in advanced cancer. *Cancer.* 1976;38:388–394.
34. Buyse M, Thirion P, Carlson RW, Burzykowski T, Molenberghs G, Piedbois P. Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. Meta-Analysis Group in Cancer. *Lancet.* 2000;356:373–378.
35. Bruzzi P, Del Mastro L, Sormani MP, et al. Objective response to chemotherapy as a potential surrogate end point of survival in metastatic breast cancer patients. *J Clin Oncol.* 2005;23:5117–5125.
36. Ratain MJ. Phase II oncology trials: let's be positive. *Clin Cancer Res.* 2005;11:5661–5662.
37. Goffin J, Baral S, Tu D, Nomikos D, Seymour L. Objective responses in patients with malignant melanoma or renal cell cancer in early clinical studies do not predict regulatory approval. *Clin Cancer Res.* 2005;11:5928–5934.
38. Thatcher N, Chang A, Parikh P, et al. Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet.* 2005;366:1527–1537.
39. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, et al. Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med.* 2005;353:123–132.
40. Blume-Jensen P, Hunter T. Oncogenic kinase signalling. *Nature.* 2001;411:355–365.
41. Lawlor MA, Alessi DR. PKB/Akt: a key mediator of cell proliferation, survival and insulin responses? *J Cell Sci.* 2001;114:2903–2910.
42. Whiteman EL, Cho H, Birnbaum MJ. Role of Akt/protein kinase B in metabolism. *Trends Endocrinol Metab.* 2002;13:444–451.
43. Majumder PK, Febbo PG, Bikoff R, et al. mTOR inhibition reverses Akt-dependent prostate intraepithelial neoplasia through regulation of apoptotic and HIF-1-dependent pathways. *Nat Med.* 2004;10:594–601.
44. Elstrom RL, Bauer DE, Buzzai M, et al. Akt stimulates aerobic glycolysis in cancer cells. *Cancer Res.* 2004;64:3892–3899.
45. Van den Abbeele AD, Badawi RD. Use of positron emission tomography in oncology and its potential role to assess response to imatinib mesylate therapy in gastrointestinal stromal tumors (GISTs). *Eur J Cancer.* 2002;38(suppl 5):S60–S65.
46. Gayed I, Vu T, Iyer R, et al. The role of ^{18}F-FDG PET in staging and early prediction of response to therapy of recurrent gastrointestinal stromal tumors. *J Nucl Med.* 2004;45:17–21.
47. Demetri GD, von Mehren M, Blanke CD, et al. Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. *N Engl J Med.* 2002;347:472–480.
48. Stroobants S, Goeminne J, Seegers M, et al. ^{18}FDG-positron emission tomography for the early prediction of response in advanced soft tissue sarcoma treated with imatinib mesylate (Glivec). *Eur J Cancer.* 2003;39:2012–2020.
49. Cullinane C, Dorow DS, Kansara M, et al. An in vivo tumor model exploiting metabolic response as a biomarker for targeted drug development. *Cancer Res.* 2005;65:9633–9636.
50. Dorow DS, Cullinane C, Conus N, et al. Multi-tracer small animal PET imaging of the tumour response to the novel pan-Erb-B inhibitor CI-1033. *Eur J Nucl Med Mol Imaging.* 2006;33:441–452.
51. Su H, Bodenstein C, Dumont R, et al. Monitoring tumor glucose utilization by positron emission tomography for prediction of treatment response to epidermal growth factor receptor kinase inhibitors. *Clin Cancer Res.* 2006;12:5659–5667.
52. Juweid ME, Wiseman GA, Vose JM, et al. Response assessment of aggressive non-Hodgkin's lymphoma by integrated International Workshop Criteria and fluorine-18-fluorodeoxyglucose positron emission tomography. *J Clin Oncol.* 2005;23:4652–4661.
53. Urschel JD, Vasan H. A meta-analysis of randomized controlled trials that compared neoadjuvant chemoradiation and surgery to surgery alone for resectable esophageal cancer. *Am J Surg.* 2003;185:538–543.
54. Lordick F, Ott K, Weber WA, et al. Individualized indication for neoadjuvant treatment based on metabolic response assessed by 18-fluorodeoxyglucose positron emission tomography (FDG-PET) in adenocarcinoma of the esophagogastric junction (AEG) [abstract]. *J Clin Oncol.* 2005;23(suppl):318S.
55. Minn H, Zasadny KR, Quint LE, Wahl RL. Lung cancer: reproducibility of quantitative measurements for evaluating 2-[F-18]-fluoro-2-deoxy-D-glucose uptake at PET. *Radiology.* 1995;196:167–173.
56. Weber WA, Ziegler SI, Thodtmann R, Hanauske AR, Schwaiger M. Reproducibility of metabolic measurements in malignant tumors using FDG PET. *J Nucl Med.* 1999;40:1771–1777.
57. Shankar LK, Hoffman JM, Bacharach S, et al. Consensus recommendations for the use of ^{18}F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute trials. *J Nucl Med.* 2006;47:1059–1066.
58. Yamane T, Daimaru O, Ito S, Yoshiya K, Nagata T, Uchida H. Decreased ^{18}F-FDG uptake 1 day after initiation of chemotherapy for malignant lymphomas. *J Nucl Med.* 2004;45:1838–1842.
59. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med.* 1989;8:431–440.
60. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med.* 1996;125:605–613.
61. Harris ST, Watts NB, Genant HK, et al. Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis: a randomized controlled trial. Vertebral Efficacy with Risedronate Therapy (VERT) Study Group. *JAMA.* 1999;282:1344–1352.
62. Riggs BL, Hodgson SF, O'Fallon WM, et al. Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *N Engl J Med.* 1990;322:802–809.