Perspective

Progress and Promise of FDG-PET Imaging for Cancer Patient Management and Oncologic Drug Development

Gary J. Kelloff,1 John M. Hoffman,1 Bruce Johnson,2 Howard I. Scher,3 Barry A. Siegel,4 Edward Y. Cheng,5 Bruce D. Cheson,6 Joyce O’Shaughnessy,7 Kathryn Z. Guyton,8 David A. Mankoff,9 Lalitha Shankar,1 Steven M. Larson,3 Caroline C. Sigman,8 Richard L. Schilsky,10 and Daniel C. Sullivan1

Abstract

2-[18F] Fluoro-2-deoxyglucose positron emission tomography (FDG-PET) assesses a fundamental property of neoplasia, the Warburg effect. This molecular imaging technique offers a complementary approach to anatomic imaging that is more sensitive and specific in certain cancers. FDG-PET has been widely applied in oncology primarily as a staging and restaging tool that can guide patient care. However, because it accurately detects recurrent or residual disease, FDG-PET also has significant potential for assessing therapy response. In this regard, it can improve patient management by identifying responders early, before tumor size is reduced; nonresponders could discontinue futile therapy. Moreover, a reduction in the FDG-PET signal within days or weeks of initiating therapy (e.g., in lymphoma, non–small cell lung, and esophageal cancer) significantly correlates with prolonged survival and other clinical end points now used in drug approvals. These findings suggest that FDG-PET could facilitate drug development as an early surrogate of clinical benefit. This article reviews the scientific basis of FDG-PET and its development and application as a valuable oncology imaging tool. Its potential to facilitate drug development in seven oncologic settings (lung, lymphoma, breast, prostate, sarcoma, colorectal, and ovary) is addressed. Recommendations include initial validation against approved therapies, retrospective analyses to define the magnitude of change indicative of response, further prospective validation as a surrogate of clinical benefit, and application as a phase II/III trial end point to accelerate evaluation and approval of novel regimens and therapies.

FDG-PET (2-[18F]fluoro-2-deoxyglucose positron emission tomography) is an accepted and widely used clinical imaging tool in oncology. U.S. Medicare reimbursement of FDG-PET recently expanded to encompass all cancer patients participating in certain prospective studies or registries in addition to more general coverage in 10 defined oncologic settings. Primarily covered are disease diagnosis, staging, and restaging, but FDG-PET is also approved for monitoring response to therapy in locally advanced and metastatic breast cancers when a change in therapy is anticipated. Clinical trials in breast cancer and other settings [e.g., non–small cell lung cancer (NSCLC) and esophageal cancer] have shown that FDG-PET imaging can provide an early indication of therapeutic response that is well correlated with clinical outcome. FDG-PET thus has the potential to improve patient management, particularly by signaling the need for early therapeutic changes in nonresponders, thereby obviating the side effects and costs of ineffective treatment. As an early surrogate for clinical benefit, the modality also has the potential to facilitate oncologic drug development by shortening phase II trials and detecting clinical benefit earlier in phase III investigations. Studies to further explore and validate these approaches are needed and can be conducted in parallel with those employing end points now used for oncologic drug approvals.

FDG-PET is based on the reliance of tumor cells on glycolysis for energy even under aerobic conditions. The sections that follow address:

- the cellular and molecular biology of neoplasia pertaining to glucose metabolism, hypoxia, and the Warburg effect;
- the development of FDG-PET as an imaging technique;
- the application of FDG-PET in cancer diagnosis, staging, and restaging;
- the rationale for using FDG-PET to assess the response to cytotoxic as well as molecularly targeted therapeutics;
- the existing clinical data on use of FDG-PET to measure therapeutic response;
• the development path for FDG-PET as a surrogate of clinical benefit, and its value in oncologic drug development, in seven case studies.
• a summary of the development and current utility of FDG-PET and recommendations for further evaluation and validation in oncologic drug development and patient management.

Cellular and Molecular Biology of Neoplasia: Glucose Metabolism and the Warburg Effect

In the early 1920s, Otto Warburg et al. observed that cancer cells exhibit an increased rate of glycolysis (1). In most living cells, oxidative phosphorylation predominates over simple glycolysis for energy production in the presence of oxygen. Tumor hypoxia may drive, at least in part, the metabolic switch from oxidative phosphorylation to glycolysis as a means of energy generation (2). However, as observed by Warburg et al., cancer cells use glycolysis for energy production regardless of the availability of oxygen. Indeed, not all cancers with high glycolysis are hypoxic (3). Thus, the effect does not necessarily arise as a consequence of hypoxia and may independently provide a growth advantage to cancer cells. This can occur because glycolysis produces energy much faster than oxidative phosphorylation despite the loss in efficiency (glycolysis yields only 2 versus 32 mol ATP per mol glucose). By exploiting rate over yield, cancer cells can more effectively compete for limited fuel resources (4). Normal cells share the consequences of more rapid resource utilization without benefiting from the higher ATP production rate. The growth advantage of cancer cells is yet more pronounced in the oxygen-poor conditions that exist in many solid tumors, which render oxidative phosphorylation less efficient (5).

One of the key alterations associated with the high glycolytic rate of cancer cells is increased cellular glucose uptake. Facilitative hexose uptake is mediated by transmembrane transporters, termed GLUT-1-5. GLUT-1, in particular, is highly expressed in several cancers (6), including breast, NSCLC, thyroid, head and neck, colon, and esophagus. Some studies suggest a correlation with tumor grade and prognosis. Other prominent changes include increased expression of hexokinases (predominantly HK-1 and HK-2), which catalyze the first phosphorylation step in glycolysis (7). In addition to their up-regulation, up to 80% of HK-1 and HK-2 are redistributed in cancer cells (other than brain) to the outer mitochondrial membrane, where the enzymes are bound via a NH2-terminal hydrophobic tail (8). This binding is thought to provide access to intramitochondrial ATP stores, limit inhibition by glucose-6-phosphate, and improve coordination among glycolysis, oxidation of glucose to lactate, and protein synthesis (7). Also modulated during carcinogenesis are other enzymes comprising the glycolytic pathway (e.g., aldolase and enolase; ref. 9) as well as regulators of glycolytic flux (e.g., 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase; ref. 10). A related change is the activation of the hexose monophosphate shunt through which glucose provides the carbon skeletons to meet the high needs of tumor cells for biosynthesis of nucleic acids and other molecules (11).

Oncogenic signal transduction pathways seem to directly stimulate glycolysis. The genes encoding glucose transporters and hexokinases are up-regulated on transfection with oncogenes, including src, ras, and c-myc, and stimulation with growth factors (12–15) independent of hypoxia. Although the molecular regulatory mechanisms are not fully elucidated, the multifunctional protein kinase Akt is postulated to play a central role in transducing these signals to metabolic as well as cell survival and proliferation pathways. Akt is activated by phosphatidylinositol 3-kinase (PI3 kinase), is negatively regulated by the dual-specificity phosphatase and tensin homologue (PTEN), and phosphorylates mammalian target of rapamycin (mTOR). PI3 kinase–dependent Akt stimulation regulates glucose metabolism in response to certain growth factor stimuli (16). In addition, activated Akt can maintain the mitochondrial membrane potential and induce hexokinase activity in cultured leukemic cells (17). Constitutive activation of Akt in human glioblastoma cells was sufficient to stimulate glucose uptake and aerobic glycolysis independent of influencing proliferation; a PI3 kinase inhibitor blocked the effect (18). Thus, Akt seems to be a key mediator of the establishment and maintenance of glycolysis in cancer cells.

Although it can also be controlled by other signaling pathways, the hypoxia-inducible factor-1α (HIF-1α) is one downstream mediator of Akt that contributes to the regulation of glycolysis. A subunit of the basic-helix-loop-helix transcription factor HIF-1, HIF-1α is a known regulator of more than a dozen genes involved in glucose transport and metabolism (9, 19). Well known to mediate the response to hypoxia, the translation and stability of HIF-1α and other transcription factors can also be stimulated under normoxic conditions by growth factors, cytokines, and other oncogenic signals (e.g., activating ras or src mutations or p53; von Hippel Lindau, or PTEN loss) via the PI3 kinase/Akt and mitogen-activated protein kinase pathways (20–23). For example, under normoxic conditions, growth factor–mediated synthesis of HIF-1α can be blocked by rapamycin (24) and mTOR overexpression stabilizes and transactivates HIF-1α (25). In a mouse prostate model overexpressing human Akt (26), mTOR inhibition reversed the neoplastic phenotype and blocked up-regulation of HIF1α target genes (including glycolytic enzymes; ref. 27). Interestingly, HIF-1α did not play a role in the stimulation of aerobic glycolysis by Akt observed in the recent study by Elstrom et al. (18), suggesting a role for other effectors of Akt. Indeed, multiple and perhaps redundant signaling molecules may control distinct steps in the activation of genes controlling glycolysis (23).

The History and Science of FDG-PET Development as an Imaging Probe

Development of 2-[18F]fluoro-2-deoxyglucose imaging. Deoxyglucose was initially developed as a chemotherapeutic agent to block the accelerated glycolysis of tumor cells, but its central nervous system toxicity was prohibitive. The current technique for FDG assessments (28, 29) was adapted from the [14C]deoxyglucose method for measuring local glucose utilization in the brain described by Sokoloff et al. (30). FDG initially follows the same metabolic pathway as glucose. Like glucose, 2-deoxyglucose is carried into the cell by glucose transporters, where it is phosphorylated by hexokinase to yield 2-deoxyglucose-6-phosphate. Whereas glucose-6-phosphate subsequently
undergoes isomerization to fructose-6-phosphate, further catabolism of 2-deoxyglucose-6-phosphate is not possible because it lacks an oxygen atom at the C-2 position. As 2-deoxyglucose-6-phosphate is unable to diffuse out of cells and the dephosphorylation reaction occurs slowly, it becomes trapped and in fact accumulates at a rate proportional to glucose utilization (Fig. 1). Because of these properties, FDG can be exploited to assess glucose uptake and metabolism.

The advantages of using $^{18}\text{F}$ include its ability to be incorporated into molecules of interest in place of hydrogen atoms or hydroxyl groups without appreciably affecting biological function. In addition, its relatively long half-life (110 minutes) obviates the need for a cyclotron at the PET facility and allows for commercial distribution of $^{18}\text{F}$-labeled radiopharmaceuticals. As a positron-emitting radionuclide such as $^{18}\text{F}$ decays, a positron is ejected from the nucleus and scatters. When its kinetic energy is dispersed, the positron combines with an electron and the two particles are then annihilated. Their rest mass is converted to two 511-keV photons emitted 180 degrees apart. If the two photons are detected in coincidence by a pair of detectors, the annihilation event can be localized along a straight line joining the coincidence detector pair. Mathematical reconstruction methods, corrected for photon attenuation and scatter, can estimate the location and quantity of positron-emitting radionuclides within an object. Whole-body PET scanning methodology was developed in the late 1980s, providing opportunity for oncologic applications of FDG-PET (31). PET and other medical imaging modalities have also been applied in preclinical research as an aid to basic research endeavors and to bridge to eventual clinical studies (32).

**Analyses of 2-[$^{18}\text{F}$]fluoro-2-deoxyglucose data.** The utility of FDG is further enhanced by the high specific activity of the labeled compound and the sensitivity of the PET scanner(s). These features allow injection of a tracer dose (e.g., nanomole), so that the underlying biological processes remain undisturbed. Thus, clinical FDG-PET studies are conducted as tracer kinetic experiments. The most accurate method to analyze these data is to quantitatively assess the FDG uptake rate over time, for example, by using kinetic modeling together with nonlinear regression techniques (33). The metabolic rate for glucose is calculated from the time course of radiotracer concentration in tissue and in arterial blood. Although metabolic rate for glucose is not dependent on uptake time, dynamic scanning following injection as well as an arterial input function are required. For thoracic studies, the latter can be derived from vascular structures within the field of view, but arterial catheterization and frequent sampling are otherwise usually required. The method enables evaluation of GLUT and hexokinase activity and accounts for dephosphorylation. The primary limitations are that tissue compartments are assumed to be homogeneous and that nonlinear regression is sensitive to noise leading to less accurate results for smaller regions. Metabolic rate for glucose is calculated based on the assumption that the lumped constant is known and does not change over time. The lumped constant describes differences in transport and phosphorylation between glucose and FDG in a specific tissue or tumor type. The metabolic rate for glucose is thus estimated according to the following formula:

$$\text{MR}_{\text{glc}} + \frac{C_{\text{glc}}}{\text{LC}} \times \frac{K_1 k_3}{(k_2 + k_3)} = \frac{C_{\text{glc}}}{\text{LC}} \times K_i$$

where $\text{MR}_{\text{glc}}$ is the metabolic rate for glucose; $C_{\text{glc}}$ is the circulating glucose level; LC is the lumped constant; $K_1$ and $k_2$ are the forward and reverse rate constants for FDG capillary transport, respectively; $k_3$ is the FDG phosphorylation rate constant; and $K_i$ is the net rate of FDG influx (see also Fig. 1). One simplified quantitative technique is the linearized Patlak analysis, which still requires dynamic scanning but fewer frames (34, 35). Patlak analysis essentially simplifies the computation of the influx rate constant, $K_i$, assuming equilibrium between FDG in tissue and plasma and negligible

---

**Fig. 1.** Transport and metabolism of FDG. Following facilitated transport by GLUT-1 ($K_1$), FDG is phosphorylated by hexokinase ($k_3$). FDG-6-phosphate can neither undergo further metabolism nor diffuse out of cells. As the dephosphorylation ($k_4$) reaction also occurs slowly, FDG-6-phosphate is trapped intracellularly and accumulates.
dephosphorylation. Importantly, the method is less sensitive to noise and it is thus possible to perform calculations at the pixel level. Automation is possible, reducing observer variability. Other quantitative methods include the two regions of interest/six-variable model, correlative imaging, the total lesion evaluation method, and the simplified kinetic method (reviewed in ref. 33).

Other methodologies for analyzing these data include visual and semiquantitative evaluations of the accumulated FDG since the net accumulation is proportional to the rate of glycolysis. These approaches assume that FDG uptake is virtually complete and that dephosphorylation of FDG-6-phosphate is negligible (see Fig. 1). The simplest approach is subjective, qualitative visual evaluation of static images (corrected for attenuation) or of whole-body images acquired at multiple axial positions. Alternatively, the radiotracer concentration can be estimated from attenuation-corrected images of the region of interest with semiquantitative analyses.

The standardized uptake value (SUV) is a semiquantitative index of tumor uptake normalized to the injected dose and some measure of the total volume of distribution, such as the patient's body weight. The following formula is one example of how SUV can be calculated:

$$\text{SUV}_{\text{body weight}} = \frac{\text{tracer concentration (nCi/g)}}{\text{injected dose (nCi)/body weight (g)}}$$

The SUV is dependent on patient size, time between injection and scan (uptake period, usually 60 minutes), plasma glucose levels, and method of image reconstruction (36–38). Normalizing SUV to body surface area or lean body mass reduces dependency on body weight, which can decrease during cancer therapy (39). It may also be appropriate to use lean body mass in heavier patients with a higher fraction of total body fat. Although the SUV can be normalized to the blood glucose concentration if it is expected to change with treatment, one study found no improvement in reproducibility between scans in treated cancer patients with correction for blood glucose (40).

SUVs are strongly correlated with the FDG metabolic rate, particularly when body surface area rather than body weight is used to calculate SUV. For example, Minn et al. found a highly significant correlation ($P < 0.0001$) between FDG metabolic rate and SUV adjusted for injected dose and weight ($r = 0.91$) or dose and body surface area ($r = 0.94$). Absolute SUV values also correlated with Patlak slope in a comparative study of 13 patients ($r = 0.97$, $P < 0.0001$), although differences across serial scans were noted for the two methods (41). Nonetheless, serial assessments have established the reproducibility of FDG-PET scans, with several studies finding an ~10% variability (refs. 40, 42; reviewed in ref. 33). Thus, regardless of analytic technique, the assessment of FDG in clinical oncology applications has been proven reliable and robust. Because it does not require dynamic data acquisition or arterial blood sampling, the SUV has frequently been used as a measure of FDG uptake to assess differences between scans.

Weaknesses of FDG-PET for cancer imaging include its limited reconstructed spatial resolution of 4 to 10 mm in commercially available scanners. Therefore, a negative scan cannot exclude the presence of a small tumor, and precise anatomic localization of the signal can be problematic in some settings (e.g., in the head and neck). Dual PET/computed tomography (PET/CT) scanners can often resolve indeterminate findings on FDG-PET alone. FDG-PET also cannot distinguish diseases of different histologic origin (e.g., carcinoma versus lymphoma of the breast) because glycolysis is a general property of malignancy. Some tumors (e.g., mucinous carcinomas and most prostate carcinomas) have relatively low FDG uptake and may not be detected by FDG-PET. In addition, nonspecific signals can arise from normal glucose uptake (e.g., in the brain or bowel) as well as inflammation and other conditions. A dynamic imaging approach, or two-point or delayed imaging, may discriminate cancer from inflammation, because FDG uptake in inflammatory sites is initially rapid and then tapers gradually after ~60 minutes, whereas tumor FDG uptake continues to increase with time. In a retrospective study of 76 patients with either malignant or benign conditions, the SUVs of malignant lesions (lung cancer, mesothelioma, non-Hodgkin's lymphoma, and esophageal cancer) increased (mean, 19.2%), whereas those of benign processes (pulmonary nodules, tubercular lesions, radiation-induced inflammation, and periprosthetic infection) tended to decrease (mean, ~6.3%) over the interval between scans (mean time, 52 minutes; range, 41–65 minutes; ref. 43). Nonspecific FDG-PET signals can also arise from fat (e.g., brown fat depots in the neck and pericardial fat) or dense muscle activity (e.g., head and neck muscles and diaphragm), but these can usually be distinguished from true malignancies by combining conventional imaging technologies with FDG-PET (e.g., PET/CT), an increasingly common approach (44). Finally, variability can be also be reduced by using patients as their own controls when quantifying differences across serial scans (e.g., in response monitoring). As detailed in the section below, additional, cancer type–specific potential sources of false-positive and false-negative FDG-PET signals—and approaches for minimizing or avoiding them—should be considered during planning and interpretation of FDG-PET scans.

FDG-PET Validated for Staging and Diagnosis: A Biomarker of Prognosis and Progression

FDG-PET has been used in cancer patients for ~25 years. In their review of the FDG-PET oncology literature from 1993 to 2000, Gambhir et al. estimated an overall average sensitivity of 84% (based on >18,402 patient studies) and a specificity of 88% (based on 14,264 patient studies) in cancer (45). As reviewed in detail by Gambhir et al. (45) and others, hundreds of publications have established a role for FDG-PET in the clinical management of cancer patients and as a biomarker of disease prognosis and progression. The Centers for Medicare and Medicaid Services (CMS) have now approved Medicare reimbursement for FDG-PET imaging in 10 oncologic settings (Table 1); in all other cancers, coverage applies only to FDG-PET scans conducted in certain prospective trials or patient registries. In most cancers, FDG-PET is approved for use in disease diagnosis, staging, and restaging. The modality is covered for diagnosis in those settings where FDG-PET results can replace invasive diagnostic procedures or can inform the anatomic location of procedures to be done. Because a tissue diagnosis is made before FDG-PET scanning in most solid tumors, FDG-PET is preferentially used to stage (rather than diagnose) melanoma, lymphoma, esophageal, and colorectal cancers. In these and other diseases, FDG-PET has had a significant impact because accurate staging is essential for
appropriate clinical management of the identified cancer. One application where FDG-PET has found particular utility is in detecting distant metastases (e.g., in melanoma and colorectal cancer) and metastatic disease in lymph nodes that appear normal on CT scan (e.g., in lymphoma, lung, and esophageal cancer; ref. 46). After treatment, FDG-PET is valuable for restaging and is used in this setting to detect recurrent or residual disease or to determine the extent of a known recurrence. In locally advanced and metastatic breast cancer, FDG-PET is also employed to monitor therapy response when a change in therapy is anticipated. The following sections address the utility of FDG-PET in CMS-approved settings. Several other cancers, including those recently considered by CMS for approval, are also presented.

### Lung (solitary pulmonary nodules and non–small cell lung cancer)
Characterization of solitary pulmonary nodules was among the first nonneurologic or noncardiac uses of FDG-PET. A considerable volume of published data support the utility and accuracy of FDG-PET in assessing and differentiating malignant from benign pulmonary nodules. For example, in a meta-analysis comprising 1,474 nodules, Gould et al. reported an overall sensitivity of 97% and a specificity of 78% for FDG-PET (47). Although well-differentiated adenocarcinoma, bronchoalveolar cell carcinoma, and low-grade neuroendocrine carcinoma can yield false-negative results, and inflammation or infection (e.g., tuberculosis and histoplasmosis) can generate false-positive signals, most hypermetabolic lesions (i.e., with SUV ≥2.5) are malignant (48). FDG-PET is primarily used following radiographic identification of the lung nodule or mass, particularly when CT findings are indeterminate. A positive diagnosis typically requires only a single FDG-PET scan, whereas serial follow-up over 3 to 6 months is usual with CT (49). However, CT follow-up to indicate stability or resolution of benign lesions is necessary due to the 3% false-negative rate of FDG-PET.

Evaluation of NSCLC is one of the primary clinical applications of FDG-PET. FDG-PET is widely used in diagnosing, staging, defining the treatment plan, and assessing recurrence for the disease (see refs. 50–53 for reviews). Several studies have found FDG-PET to be more accurate than conventional imaging in lung cancer. For example, in 100 lung cancer patients, the accuracy of FDG-PET for staging was 83% versus 65% by chest CT and bone scintigraphy (54). In their meta-analysis, Tozola et al. found FDG-PET to be 84% sensitive and 89% specific; the modality was more accurate than CT or endoscopic ultrasound for staging the mediastinum (55). FDG-PET is also superior to CT and magnetic resonance imaging (MRI) for detecting metastases, except for metastatic brain lesions where FDG-PET is clearly less sensitive than contrast-enhanced CT and MRI (52, 54) due to normal high rate of glucose metabolism in the brain. Compared with CT, FDG-PET improved detection of local and distant metastases, altering the clinical stage in 62 of 102 NSCLC patients studied (56). In a separate study, FDG-PET corrected the clinical stage in 27% and detected metastases in 13% of 97 patients (57). Similarly, FDG-PET upstaged 30% of 57 patients studied and improved selection for combined modality treatment by eliminating those with metastatic disease after induction therapy (58). Across series, a change in patient management occurred in up to half or more of patients (59). For example, Kalff et al. found that FDG-PET altered or influenced management in 70 of 105 (67%) patients studied (60). In the randomized, controlled PLISS trial, 39 (41%) of patients undergoing conventional imaging had thoracotomy that was determined to be futile compared with 19 (21%) in the patient group who also had FDG-PET imaging (61). The American

### Table 1. CMS coverage for FDG-PET in oncology

<table>
<thead>
<tr>
<th>Clinical condition</th>
<th>Coverage (effective date)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solitary pulmonary nodule</td>
<td>Characterization (January 1998)</td>
</tr>
<tr>
<td>Lung cancer (NSCLC)</td>
<td>Initial staging (January 1998)</td>
</tr>
<tr>
<td>Esophageal cancer</td>
<td>Diagnosis, staging, and restaging (July 2001)</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Diagnosis, staging, and restaging (July 2001)</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Tumor localization if carcinoembryonic antigen suggests recurrence (July 1999)</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Diagnosis, staging, and restaging (July 2001)</td>
</tr>
<tr>
<td>Head and neck cancer (excluding central nervous system and thyroid)</td>
<td>Diagnosis, staging, and restaging (July 2001)</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>Adjunct for diagnosis, staging, restaging, and monitoring response (October 2002), particularly as an adjunct for staging metastatic disease, restaging locoregional recurrence or metastasis, and monitoring response of locally advanced and metastatic breast cancer when a change in therapy is anticipated</td>
</tr>
<tr>
<td>Thyroid cancer (follicular cell)</td>
<td>Evaluating recurrent or residual follicular cell tumors (treated previously by thyroidectomy and radioiodine ablation) when serum thyroglobulin &gt;10 ng/mL and 131I whole-body scan is negative (October 2003)</td>
</tr>
<tr>
<td>Cervical cancer</td>
<td>Detecting pretreatment metastases in newly diagnosed cervical cancer after negative conventional imaging (January 2005)</td>
</tr>
</tbody>
</table>
College of Surgeons Oncology Group Z0050 trial also found that one in five patients could avoid unnecessary surgery based on FDG-PET data (62). The significant reduction in planned surgery with FDG-PET supports the cost-effectiveness and benefit to the patient of the approach. However, in a separate randomized, controlled trial in which advanced-stage cancer was rarely identified (2 cases among the 184 enrolled stage I-II patients), FDG-PET did not reduce unnecessary thoracotomies but did influence patient management (63). The prognostic significance of FDG-PET in NSCLC has also been established (64, 65). For example, FDG-PET-detected metastatic tumor burden was correlated with survival in a study of 42 NSCLC patients (66). Interestingly, a recent study of 178 NSCLC patients suggested an alternative basis, besides tumor grade, of the prognostic value of FDG-PET (67). Partial volume correction (for tumor size) abolished the significant correlation between SUV<sub>max</sub> and surgical tumor stage; a longer follow-up is planned to assess the prognostic significance of the partial volume-corrected FDG-PET signal.

FDG-PET also has utility for restaging both local and metastatic recurrences of NSCLC. As in initial NSCLC staging, FDG-PET is highly specific for detecting metastatic disease (e.g., to lymph nodes, liver, bone, and adrenal glands). FDG-PET also offers improved sensitivity over CT imaging for differentiating new pulmonary nodules from scar tissue arising after surgical resection, radiation, and chemotherapy. For example, compared with CT, FDG-PET was 100% versus 71% sensitive and 92% versus 95% specific, respectively, in 126 NSCLC patients assessed before and after therapy (68). In a separate study of 63 NSCLC patients with suspected relapse, FDG-PET was 98% sensitive and had a negative predictive value of 93%: FDG-PET results stimulated a change in management of 40 (63%) patients (69). Similarly, in 156 NSCLC patients initially evaluated by CT and referred for restaging, FDG-PET downstaged 29% and upstaged 33%, with a resulting reclassification (from resectable to unresectable or vice versa) of 37% of patients (70). These and numerous other studies and clinical trials (see ref. 45 for a comprehensive review) provide the basis for the 1998 and 2001 CMS approvals for NSCLC.

Esophagus. Although relatively uncommon in the United States, esophageal cancer is associated with high mortality and thus accounts for >11,000 cancer deaths per year (71). FDG-PET can identify known primary esophageal tumors but lacks accuracy for regional nodal disease because of proximity to the primary lesion and the often microscopic nature of the neoplastic foci. In one study of 42 patients, FDG-PET was insensitive for regional nodes but superior to combined assessment with ultrasound and CT imaging for evaluating distant nodal metastases (72). In their recent meta-analysis, van Westreenen et al. found an overall sensitivity of 51% and specificity of 84% for locoregional metastases compared with a pooled sensitivity and specificity of 67% and 97%, respectively, for distant metastases (73). Because of its accuracy in identifying and characterizing metastatic disease, FDG-PET is primarily used in esophageal cancer to stage and plan treatment in patients being considered for resection (74). A combined modality approach (i.e., PET + CT) has been advocated as the most accurate method for staging such patients, clarifying clinical management decisions in 90% of 26 cases in one study (75); recent data from Bar-Shalom et al. support a greater accuracy of combined modality PET/CT imaging than FDG-PET alone or separately conducted FDG-PET and CT scans (76). FDG also has prognostic value in esophageal cancer. In a retrospective analysis of 32 patients, FDG uptake was significantly associated with depth of tumor invasion, presence of lymph node metastasis, and lymphatic invasion. Moreover, high FDG uptake in the primary tumor SUV (>3) significantly correlated with lower survival (77). Esophageal cancers commonly recur. Although FDG-PET is sensitive for local recurrence, specificity is limited by uptake due to inflammation, benign disease, and other conditions (e.g., following balloon dilation). However, FDG-PET is superior to conventional imaging for disease that recurs outside of the surgical field (75).

Head and neck. Outside of Asia, head and neck cancers are uncommon and comprise only 2% to 4% of U.S. cancers. Of these, most are oral and laryngeal cancers, which are accessible for diagnosis by visual and physical examination; FDG-PET evaluation does not usually provide additional diagnostic information. However, FDG-PET does have utility for initial diagnosis in those patients presenting with confirmed metastases in the cervical lymph nodes but unknown primary tumor. In studies comprising nearly 300 patients, FDG-PET determined the location of primary disease in 10% to 60% of cases (reviewed in refs. 49, 78). In addition, FDG-PET can define the extent of the primary disease before and after chemoradiotherapy; in contrast to estimates based on conventional imaging, the FDG-PET-defined extent of disease was a significant predictor of survival ($P < 0.0001$) in one recent study (79). FDG-PET is particularly accurate for staging local nodal spread, a key factor for prognosis and treatment planning. Although they can be readily detected by FDG-PET, distant metastases in head and neck cancer are uncommon, and second primaries are estimated to occur in 8% of cases (78, 80). In detection of recurrent disease, FDG-PET has greater sensitivity and specificity than conventional imaging; CT is limited by the anatomic distortion commonly seen following treatment due to inflammation and edema and typically requires serial examinations. In 53 patients with residual structural abnormalities following definitive treatment, FDG-PET changed patient management in 40%; planned surgery was determined to be futile in 14 patients based on negative FDG-PET scan (79). Indeed, FDG-PET has a high negative predictive value (89%) in one study of 75 patients; ref. 81), whereas positive results are less reliable. Because FDG-PET abnormalities are also imprecisely localized, combined modality imaging (PET/CT) is advocated (82).

Colorectum. In colorectal cancer, the primary utility of FDG-PET is in combination with standard CT imaging to detect distant metastases. The modality has low specificity (40-60%) for colorectal cancer because FDG accumulation occurs physiologically in the bowel wall and is enhanced when inflammation and colon polyps are present. FDG-PET also has limited utility for local and regional staging, with sensitivity for regional lymph node involvement of only ~29% (83). Combined modality PET/CT offers improved specificity for the primary neoplasm (84). FDG-PET is a valuable addition to CT imaging for characterizing hepatic metastases (particularly those >1 cm) and for detecting extrahepatic metastases.
In their meta-analysis, Huebner et al. found an overall sensitivity and specificity of 97% and 76%, respectively, of FDG-PET for detecting colorectal metastases throughout the body (90). FDG-PET is thus a particularly important staging tool in patients with metastatic disease considered for curative hepatic resection (91, 92). FDG-PET detected additional extrapathologic disease in ~25% of 43 patients studied and thereby disqualified them for hepatic resection (93); survival after surgery was increased in this and a later study of 100 patients because of the improved selection of patients for the procedure (93, 94).

Melanoma. Melanoma is another setting in which FDG-PET is highly sensitive and specific for metastatic disease and is therefore an important tool for surgical planning (95). FDG-PET is inferior to sentinel lymph node mapping for characterizing spread to regional lymph nodes because only microscopic disease is present in the sentinel node in most cases (96, 97). However, malignant melanoma can spread to unusual and various sites (e.g., gallbladder, adrenal glands, and bone) that can be easily missed by conventional imaging (98). FDG-PET thus has particular utility for evaluating patients in whom distant metastases are suspected and is used in surveillance of high-risk stage III and IV patients after treatment (99). Post-treatment FDG-PET imaging is especially useful because aggressive resection of metastatic loci is a typical surgical approach. In one study, FDG-PET altered the therapeutic plan in 90% of 34 enrolled patients (100).

Lymphoma. Lymphomas comprise ~30 distinct diseases, which are broadly divided into Hodgkin’s disease and non–Hodgkin’s lymphoma types. Classification (using the REAL or WHO systems) based on morphology, cell surface markers, genetic abnormalities, and clinical features of the disease is essential for guiding treatment and anticipating outcome. Accurate staging also directs treatment selection and thereby improves outcome. FDG-PET is not used for diagnosis in lymphoma because excisional lymph node biopsy with histopathology and immunohistochemistry and flow cytometry is the standard. Small series suggest that FDG uptake is correlated with tumor grade at biopsy (e.g., ref. 101), but discordant results may occur (e.g., in nodal large-cell non–Hodgkin’s lymphoma with follicular involvement in the bone marrow). Assessing bone marrow involvement by biopsy is also an essential part of patient evaluation and was superior to FDG-PET in a retrospective study of 172 patients (102). However, noninvasive imaging with FDG-PET can play an important role in staging, and several studies have shown its superiority to anatomic imaging modalities (e.g., refs. 103–106; reviewed in refs. 107, 108). For example, compared with staging by CT, FDG-PET was equally sensitive but significantly more specific for Hodgkin’s disease (96% versus 41%) as well as non–Hodgkin’s lymphoma (100% versus 67%) in a retrospective study of 50 patients (104). Similarly, FDG-PET gave a lower (28%) or higher (12%) stage in 81 Hodgkin’s disease patients (106). FDG-PET is also superior to, and has largely replaced, 67Ga scintigraphy for staging (109). Hypermetabolic conditions (sarcoidosis, tuberculosis, fungal infections, etc.) are a source of false-positive findings in lymphoma; low-grade tumors and certain lymphomas (e.g., peripheral T-cell and marginal zone, including mucosal-associated lymphoid tissue) have low FDG uptake that can result in false-negative scans (102). Conventional imaging has low specificity for distinguishing residual tumor from fibrosis or scar tissue following therapy (104). Thus, FDG-PET is being increasingly used to restage lymphoma and has particular utility in assessing malignancy in residual masses post-treatment.

Breast. Accurate staging and restaging is essential for optimal management of invasive breast cancer. FDG-PET has utility for initial staging, defining the extent of disease, and treatment planning, particularly for patients with recurrent or metastatic disease. FDG-PET is approved for these uses as an adjunct to conventional imaging approaches. In small tumors and certain low-grade cancers (e.g., tubular and lobular carcinomas and ductal carcinoma in situ), limited FDG accumulation can cause false-negative results, whereas false-positive results can arise due to inflammation (49, 110). FDG-PET is specific (79–100%) for detecting axillary nodal disease, but the sensitivity of FDG-PET is low in cases when the involved nodes or metastases are small (~5 mm; refs. 111–114). Sentinel lymph node mapping is superior overall; although FDG-PET has the advantage of being noninvasive (115), most cancers present as stage I or II disease with no or small volume disease in the axilla and thus the clinical utility of FDG-PET as an axillary staging tool is low. Nonetheless, FDG-PET is 2-fold more sensitive than CT for mediastinal or internal mammary nodes and is helpful in planning treatment (e.g., nodal radiation) for advanced axillary disease (116, 117). Further, in settings where breast cancer has metastasized beyond the axillary lymph nodes, FDG-PET has equal (bone and lung) or superior (liver) specificity and sensitivity relative to CT (118–122). As for restaging of other cancers, FDG-PET is useful for differentiating locally recurrent disease from scar tissue and fibrosis and for detecting systemic metastases following definitive treatment. FDG-PET correctly confirmed suspected recurrent or metastatic disease in 25 of 27 patients (123). In a recent retrospective study of 125 recurrent or metastatic breast cancer patients, FDG-PET had a significant impact in defining the extent of disease and, consequently, the planned therapeutic approach. The treatment plan was altered in 32% and supported in 27% of patients. A change was most likely in patients with suspected locoregional recurrence (124).

Thyroid. CMS has approved the use of FDG-PET in thyroid cancer patients with elevated serum thyroglobulin but negative 131I whole-body scan. Several studies have confirmed the high sensitivity and specificity of FDG-PET in these patients in whom FDG-PET can localize metastatic disease and thereby guide management (see refs. 49, 78 for reviews). For example, FDG-PET changed the surgical approach in 9 of 24 patients in one study (125) and in 19 of 24 in another study (126). Several studies also support the prognostic value of FDG-PET in thyroid cancer; a high volume of FDG-avid disease was associated with reduced survival in one study of 125 thyroid cancer patients followed for 41 months (127). Therefore, FDG-PET is highly useful in clinical decision-making regarding how aggressively to treat thyroid cancer, a disease that can range from rather indolent to highly aggressive. Emerging data suggest that FDG-PET may also have utility in certain rare but aggressive thyroid cancers (particularly Hürthle cell carcinoma and anaplastic thyroid cancer; reviewed in ref. 78). In addition, incidentally discovered focal FDG uptake in the thyroid gland has been associated with malignancy in up to 50% of cases, whereas diffuse uptake is indicative of thyroiditis (128, 129).
Other settings. Accumulating evidence suggests that FDG-PET is a promising imaging modality for other malignancies as well. For example, the utility of FDG-PET has been shown in grading of sarcomas (130–133). Emerging data in testicular cancer support the utility of FDG-PET in initial staging as well as for evaluating residual disease after chemotherapy or if serum markers are elevated but CT scans are negative (134–138). In cervical cancer, several investigations have shown the utility of FDG-PET for identifying occult metastases and for evaluating recurrence and therapy response (139–145). CMS recently considered Medicare approval in six additional cancer settings—small cell lung cancer as well as cancers of the brain, pancreas, cervix, ovary, and testes [see CMS Web site for detailed review of the evidence (146)]. In their January 2005 Decision Memorandum (147), CMS expanded Medicare approval to include FDG-PET scans for the detection of pretreatment metastases in newly diagnosed cervical cancer after negative conventional imaging. In addition, CMS coverage was expanded to include all cancer patients participating in a prospective clinical study of the following types:

- trial meeting requirements for Food and Drug Administration category B investigational device exemption or
- an appropriately designed and conducted FDG-PET clinical study to collect additional information at the time of the scan to assist in patient management.

Current Therapeutics for Oncologic Disease: Mechanistic Rationale for FDG-PET as a Measure of Activity

Oncologic treatment options typically include surgery and radiotherapy either alone or together with combination chemotherapy. Using gene expression arrays and other approaches, recent studies have sought to characterize the response to chemotherapeutic agents to elucidate their molecular effectors (148). As depicted in Fig. 2, emerging data suggest that cytotoxic and cytostatic agents affect, directly or indirectly, the pathways, glucose transporters, and metabolic enzymes controlling glycolysis. For example, cytotoxic agents, such as cisplatin and etoposide, dramatically down-regulate hexokinases as well as GLUT-1 and GLUT-3, and suppress glycolysis in vitro (149). Cyclophosphamide also inhibits glycolysis as shown in C3H mice with radiation-induced fibrosarcomas (150). Paclitaxel inhibits glycolysis by mediating detachment of phosphofructokinase from the cytoskeleton, resulting in decreases in two allosteric stimulators of glycolysis (151). Topotecan may affect transcription of the genes controlling glycolysis by decreasing the rate of HIF-1 protein translation (152). In human breast cancer xenografts, estrogen-stimulated growth is associated with a dramatic increase in tumor glycolytic activity and a concomitant elevation in GLUT-1 expression. Tamoxifen treatment induced growth arrest, halved the rate of glycolysis, and dramatically decreased GLUT-1 expression (153).

A correlation between efficacy and glucose metabolism has now been established for targeted therapies, such as imatinib (Gleevec; ref. 154). Other approved cytostatic agents are known to influence signaling through the mitogen-activated protein kinase and/or Akt pathways; as noted above, transfection experiments have shown that Akt, in particular, plays a direct role in stimulation and maintenance of aerobic glycolysis (18). Recent data suggest that certain epidermal growth factor receptor mutations conferring increased pathway signaling may be important determinants of gefitinib (Iressa) sensitivity (155–157). A recent study of 109 NSCLC patients treated with gefitinib associated higher response rate, longer time to progression (TTP), and a reduced risk of disease progression with Akt (but not mitogen-activated protein kinase) positivity (158). Similarly, Han et al. reported that Akt positivity significantly correlated with prolonged TTP ($P = 0.018$) and overall survival (OS; $P = 0.008$) following gefitinib treatment of 65 NSCLC patients (159). A recent in vitro study suggests that gefitinib may induce apoptosis via Akt inhibition in sensitive NSCLC cells; cells with wild-type epidermal growth factor receptor instead underwent G_{S}S growth arrest (160).

The correlation of molecular abnormalities in specific cancers with alterations in glucose metabolism, including transporters, which are concordantly modulated with several classes of chemotherapeutic drugs, suggests the utility of FDG-PET for therapy monitoring. One possible confounder is that agents targeting Akt or other molecular signals could theoretically affect glycolysis without chemotherapeutic efficacy on the disease or survival. Gene expression arrays or other assays may aid in interpretation of FDG-PET imaging data where the precise molecular target or mechanisms for affecting glycolysis are unknown. In addition, preclinical studies may help to clarify the expected FDG-PET outcome for certain therapies. Indeed, although the data are still emerging, validation of the association of the modulation of glycolysis with tumor response is actively being pursued in ever more relevant preclinical models and innovative small phase II neoadjuvant studies. Existing clinical data, reviewed in the next section, indicate the promise of FDG-PET as a measure of treatment efficacy. Further prospective studies will provide validation in many clinical cancer target organs.

Clinical Data of FDG-PET as a Measure of Treatment Efficacy with Approved Therapies and Its Role in Patient Management

Although CMS reimbursement for monitoring response to therapy is only approved in certain breast cancer settings, FDG-PET has shown encouraging results as an early predictor of tumor response, progression-free survival (PFS), and OS in a range of clinical settings. Table 2 highlights 25 studies involving nearly a thousand patients with lymphoma, lung, esophagus, head and neck, and other cancers that have correlated the FDG-PET response with clinical outcome. As a measure of treatment efficacy, FDG-PET has several potential applications that can significantly affect patient management. These include early assessment of response, so that ineffective therapy can be discontinued. This is particularly important in disease such as NSCLC, where response rates to existing toxic therapies are modest (20-40%) and expected OS is also low. An additional potential utility of FDG-PET is in monitoring the efficacy of neoadjuvant therapy. Response in the neoadjuvant setting is an important prognostic factor in many cancers. It predicts the benefit from surgery in esophageal cancer and chemosensitivity to the same breast cancer chemotherapeutic agents postsurgery. The sections below summarize the key data validating FDG-PET as a predictor of patient outcome. In many cases, SUV cut
points were determined post hoc, and these values need to be further validated in prospective studies. The utility of the imaging modality for influencing patient management in these various clinical settings is highlighted.

**Lung.** In a prospective study by Weber et al., 57 advanced NSCLC patients were studied with FDG-PET before and after the first 21-day cycle of chemotherapy (161). Chemotherapeutic regimens included vinorelbine/cisplatin, docetaxel/cisplatin, paclitaxel/carboplatin, and cisplatin/etoposide, which exhibit similar activity and effectiveness in NSCLC (162). Response was prospectively defined as >20% decrease in SUV based on evidence that this magnitude of change could be reliably measured (40, 42). The FDG-PET response at 3 weeks closely correlated with best response to therapy as assessed by Response Evaluation Criteria in Solid Tumors at 2 months (P < 0.0001). Furthermore, FDG-PET predicted survival; for metabolic responders and nonresponders, respectively, median TTP was 163 versus 54 days (P = 0.0003) and OS was 252 versus 151 days (P = 0.005). Because metabolic responders had a high probability of achieving an objective response by standard response criteria and FDG-PET response correlated well with patient outcome, FDG-PET was a surrogate for clinical benefit in this small series. In a separate study of 73 NSCLC patients, a FDG-PET scan at 4 to 12 weeks (median, 70 days) following radiotherapy or chemoradiotherapy was superior to one-time CT imaging for predicting survival. FDG-PET and CT assessments were equivalent in only 40% of patients, with a higher response category (i.e., significantly more complete responses) determined by FDG-PET in 80% of divergent cases (163). Residual masses were evident on the post-treatment CT scan in 86% of patients, complicating anatomic evaluation. Other investigations also support the utility of FDG-PET as an early surrogate for assessing therapy response in NSCLC (69, 164, 165). For example, in 29 evaluated patients with locally advanced NSCLC, the probability of pathologic response significantly correlated with the glucose metabolic rate evaluated 2 weeks after completing neoadjuvant radiotherapy with or without

![Fig. 2. Molecular targets of cytotoxic and cytostatic drugs in the pathways controlling glycolytic metabolism.](https://www.cancerres.org/content/11/8/2793/F2.large.jpg)
chemotherapy (164). In a separate study of 26 stage III NSCLC patients evaluated 2 weeks after neoadjuvant chemoradiotherapy, FDG-PET had a sensitivity and specificity of 88% and 67%, respectively, when a SUV of 3 was used to differentiate residual tumor from pathologic complete response (165). In a third study of 60 NSCLC patients with suspected relapse following definitive treatment, both the presence and the extent of relapse determined by FDG-PET significantly predicted survival (69).

Esophagus. In locally advanced esophageal cancer, histopathologic response to preoperative chemotherapy or chemoradiotherapy is one of the most important prognostic factors. Although responders may survive three to four times longer, neoadjuvant therapy does not confer an OS benefit because very few patients do respond (166, 167). Whereas responders should undergo esophagectomy, nonresponders have such a poor prognosis (median survival, 9 months) that the benefit of surgical resection is questionable. If nonresponding patients could be identified early, futile toxic therapies could be discontinued and alternatives considered. As detailed below and in Table 2, studies comprising >130 patients support the potential of FDG-PET to predict

<table>
<thead>
<tr>
<th>Tumor/treatment (n)</th>
<th>n</th>
<th>Criteria</th>
<th>End point (P)</th>
<th>FDG-PET responder</th>
<th>FDG-PET nonresponder</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lymphoma</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRT (13) or chemotherapy (41)</td>
<td>54</td>
<td>Visual</td>
<td>1-y OS (&lt;0.0001)</td>
<td>92%</td>
<td>50%</td>
<td>(190)</td>
</tr>
<tr>
<td>Anthracycline-, mitoxantrone-, or cisplatin-based chemotherapy</td>
<td>28</td>
<td>Visual</td>
<td>OS (&lt;0.0001): 1 y</td>
<td>87%</td>
<td>20%</td>
<td>(182)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 y PFS (&lt;0.0001): 1 y</td>
<td>68%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Doxorubicin-based or CVP</td>
<td>93</td>
<td>Visual</td>
<td>PFS (&lt;0.00001): median SUV 2 y</td>
<td>62%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>High-dose chemotherapy (Hodgkin’s disease: MOPP/ABV or Stanford V regimen (269); non–Hodgkin’s lymphoma: CHOP or CHVmp/BV) + SCT</td>
<td>60</td>
<td>Visual</td>
<td>2-y OS (&lt;0.00002)</td>
<td>100%</td>
<td>55%</td>
<td>(181)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFS (&lt;0.000001): median SUV 2 y</td>
<td>85%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td><strong>Induction treatment + high-dose chemotherapy with (9) or without radiotherapy (15) / + SCT</strong></td>
<td>22</td>
<td>SUV (&gt;25%)</td>
<td>Median OS (0.006)</td>
<td>&gt;60 mo</td>
<td>14 mo</td>
<td>(183)</td>
</tr>
<tr>
<td>Chemotherapy, radiotherapy, or CRT with (2) or without (54) SCT</td>
<td>56</td>
<td>Visual SUV (below mean)</td>
<td>PFS (&lt;0.0025)</td>
<td>21 mo</td>
<td>2 mo</td>
<td>(184)</td>
</tr>
<tr>
<td>Chemotherapy (11), radiotherapy (1), CRT (8), or salvage chemotherapy with (6) or without (2) SCT</td>
<td>28</td>
<td>Visual</td>
<td>DFS (0.004)</td>
<td>95%</td>
<td>40%</td>
<td>(270)</td>
</tr>
<tr>
<td>Chemotherapy (20), CRT (34), and/or SCT (7)</td>
<td>58</td>
<td>SUV (&gt;3) in residual mass</td>
<td>PFS (&lt;0.00001)</td>
<td>&gt;48 mo</td>
<td>18 mo</td>
<td>(185)</td>
</tr>
<tr>
<td><strong>CHOP, DEXA-ICE, or DBVD</strong></td>
<td>30</td>
<td>Visual, after first cycle</td>
<td>Median OS (0.006)</td>
<td>4%</td>
<td>62.5%</td>
<td>(188)</td>
</tr>
<tr>
<td><strong>High-dose chemotherapy + SCT</strong></td>
<td>16</td>
<td>Visual</td>
<td>1-y OS (0.01)</td>
<td>100%</td>
<td>55%</td>
<td>(271)</td>
</tr>
<tr>
<td><strong>DHAP-VIM ± SCT</strong></td>
<td>46</td>
<td>Visual</td>
<td>2-y PFS (0.048)</td>
<td>62%</td>
<td>32%</td>
<td>(272)</td>
</tr>
<tr>
<td><strong>Breast</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclophosphamide, doxorubicin, premarin, methotrexate, 5-FU, tamoxifen, radiotherapy</td>
<td>11</td>
<td>SUV (&gt;48%)</td>
<td>6-mo DFS (&lt;0.0001)</td>
<td>73%</td>
<td>27%</td>
<td>(214)</td>
</tr>
<tr>
<td>Gemcitabine, epirubicin, and paclitaxel or epirubicin and paclitaxel</td>
<td>9</td>
<td>SUV</td>
<td>PFS (significant)</td>
<td>15 mo</td>
<td>5 mo</td>
<td>(218)</td>
</tr>
<tr>
<td>Neoadjuvant therapy (doxorubicin or cyclophosphamide ± 5-FU; cyclophosphamide + methotrexate + 5-FU + radiotherapy; paclitaxel or docetaxel + vinorelbine; or paclitaxel + Herceptin)</td>
<td>31</td>
<td>FDG metabolic rate</td>
<td>DFS (0.09-0.05)</td>
<td>&gt;60 mo</td>
<td>~35 mo</td>
<td>(273)</td>
</tr>
</tbody>
</table>

(Continued on the following page)
histopathologic response within weeks after initiating preoperative therapy. In contrast, endoscopic ultrasonography has an accuracy of <50%, and tumor size cannot be reliably measured with conventional imaging and may not reflect response (e.g., due to edema).

In a study of 40 patients, a decline in FDG uptake 14 days after preoperative cisplatin-based chemotherapy significantly differed among responders (54%) and nonresponders (15%; ref. 168). Using a 35% cutoff (defined in receiver operator characteristic analysis), FDG-PET predicted clinical response (>50% reduction in tumor length by standard imaging 3 months post-therapy) as well as recurrence-free survival (P = 0.01) and OS (P = 0.04). Similarly, in a study of 24 patients who went on to esophagectomy, a decreased SUV (≥52%) 3 weeks after 5-fluorouracil (5-FU) and radiotherapy correlated with both histopathologic response and survival (median, 22.5 versus 8.8 months, P = 0.0001; ref. 169). A 60% reduction in SUV after induction paclitaxel/cisplatin chemotherapy (with or without radiation) predicted survival in a fourth study of 39 patients (171); 2-year disease-free survival (DFS) and OS were 63% and 89% compared with only 38% DFS and 67% OS in patients with >60% versus <60% SUV decrease (P = 0.055 and 0.089, respectively). Finally, in a recent study of 38 patients, a reduced SUV 3 to 4 weeks after 5-FU with radiation correlated with histologic response and survival.

### Table 2. FDG-PET assessment of tumor response to standard therapy in multiple cancer types and correlation with patient outcome (Cont’d)

<table>
<thead>
<tr>
<th>Tumor/treatment (n)</th>
<th>n</th>
<th>Criteria</th>
<th>End point (P)</th>
<th>FDG-PET responder</th>
<th>FDG-PET nonresponder</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorectal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-FU (10) or tegafur (15) ± radiotherapy</td>
<td>25</td>
<td>SUV (6)</td>
<td>3 y (0.04)</td>
<td>92%</td>
<td>60%</td>
<td>225</td>
</tr>
<tr>
<td>Neoadjuvant 5-FU/leucovorin and radiotherapy</td>
<td>15</td>
<td>SUV (62%)</td>
<td>OS (0.08)</td>
<td>50 mo</td>
<td>40 mo</td>
<td>226</td>
</tr>
<tr>
<td>Esophagus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neoadjuvant 5-FU and radiotherapy</td>
<td>24</td>
<td>SUV (52%)</td>
<td>OS (0.0001)</td>
<td>22.5 mo</td>
<td>8.8 mo</td>
<td>169</td>
</tr>
<tr>
<td>Neoadjuvant 5-FU, cisplatin and radiotherapy</td>
<td>36</td>
<td>Visual</td>
<td>OS (0.002)</td>
<td>16.3 mo</td>
<td>6.4 mo</td>
<td>170</td>
</tr>
<tr>
<td>Neoadjuvant paclitaxel, cisplatin ± radiotherapy</td>
<td>39</td>
<td>SUV (60%)</td>
<td>2-y OS (0.089)</td>
<td>67%</td>
<td>38%</td>
<td>171</td>
</tr>
<tr>
<td>Neoadjuvant 5-FU, leucovorin and cisplatin ± paclitaxel</td>
<td>40</td>
<td>SUV (35%)</td>
<td>OS (0.04); 2 y</td>
<td>Median</td>
<td>&gt;53 mo</td>
<td>13 mo</td>
</tr>
<tr>
<td>Neoadjuvant 5-FU, radiotherapy</td>
<td>38</td>
<td>SUV (30%)</td>
<td>OS (0.011)</td>
<td>138 mo</td>
<td>18 mo</td>
<td>172</td>
</tr>
<tr>
<td>Lung</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRT (10) or CRT [carboplatin (47) or cisplatin + either taxol (6), 5-FU (9), or VP-16 (1)]</td>
<td>63</td>
<td>Visual</td>
<td>OS (0.012)</td>
<td>22 mo</td>
<td>&lt;15 mo</td>
<td>69</td>
</tr>
<tr>
<td>Radiotherapy (10) or CRT or CRT</td>
<td>73</td>
<td>Visual</td>
<td>OS (0.0001)</td>
<td>36 mo</td>
<td>12 mo</td>
<td>163</td>
</tr>
<tr>
<td>Vinorelbine, etoposide or docetaxel + cisplatin, or paclitaxel and carboplatin</td>
<td>57</td>
<td>SUV (20%)</td>
<td>OS (0.005): median</td>
<td>252 d</td>
<td>151 d</td>
<td>161</td>
</tr>
<tr>
<td>Head/neck</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neoadjuvant radiotherapy and cisplatin (10) or CRT (10)</td>
<td>35</td>
<td>SUV/4</td>
<td>3-y OS (0.046)</td>
<td>80%</td>
<td>43%</td>
<td>274</td>
</tr>
<tr>
<td>Neoadjuvant radiotherapy (37) or CRT (10)</td>
<td>47</td>
<td>FDG metabolic rate (below median)</td>
<td>OS (0.0042)</td>
<td>72%</td>
<td>35%</td>
<td>178</td>
</tr>
<tr>
<td>Cervix</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiotherapy alone (20) or with cisplatin (132)</td>
<td>152</td>
<td>Visual</td>
<td>5-y OS (0.001)</td>
<td>92%</td>
<td>46% (0% if new FDG uptake sites)</td>
<td>143</td>
</tr>
<tr>
<td>Radiotherapy with cisplatin</td>
<td>47</td>
<td>Visual</td>
<td>OS (0.0005)</td>
<td>73%</td>
<td>0%</td>
<td>223</td>
</tr>
<tr>
<td>Sarcoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imatinib mesylate</td>
<td>21</td>
<td>SUV [EORTC (212)]</td>
<td>1-y PFS (0.00107)</td>
<td>92%</td>
<td>12%</td>
<td>208</td>
</tr>
</tbody>
</table>

Abbreviations: CRT, chemoradiotherapy; SCT, stem cell transplant; CVP, cyclophosphamide, vincristine, prednisone; MOPP/ABV, mechlorethamine, Oncovin (vincristine sulfate), procarbazine, prednisone, Adriamycin (doxorubicin), bleomycin, vinblastine; CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone; CHMVP/BV, cyclophosphamide, doxorubicin, teniposide, prednisone, bleomycin, vincristine; Induction treatment, Anthracycline-, mitoxantrone-based chemotherapy, B-ALL protocol, or DEXA-HEAM; high dose chemotherapy, endoxan or cyclophosphamide, etoposide, and carmustine; DEXA-ICE, dexamethasone, ifosfamide, cisplatin, and etoposide; DBV, doxorubicin, bleomycin, vinblastine, and dacarbazine; DHAP-VIM, dexamethasone, cytarabine, and cisplatin followed by etoposide, ifosfamide, and methotrexate; EORTC, European Organization for Research and Treatment of Cancer criteria.
survival (172). Using a –30% SUV cutoff, positive and negative predictive values were 93% and 88%, respectively. Survival was >38 versus only 18 months in patients with >30% versus <30% SUV decrease, respectively (P = 0.011).

**Head and neck.** Head and neck cancers are clinically heterogeneous, comprising multiple anatomic sites of origin with distinct natural histories and prognoses. Cure rates are low (30-50%) in locally advanced disease. In such patients (stage III-IV), the accuracy of FDG-PET for detecting residual disease provides particular utility for assessing response to neoadjuvant or definitive treatment. Following treatment, prompt salvage surgery improves local control, but the procedure can be avoided or reduced (i.e., for organ preservation) in responding patients. It may also be possible to limit or avoid neck dissection if lymph nodes are shown to lack disease following treatment. Several studies have shown that FDG-PET can assess treatment response, influence patient management, and predict histopathologic control and outcome (reviewed in ref. 173). For example, in a Japanese study, 7 of 15 patients with reduced FDG-PET 4 weeks postchemoradiotherapy (SUV <4) lacked viable tumor cells and either avoided surgery or underwent a less extensive procedure (174). In four patients, positive FDG-PET scans were due to severe mucositis; an interval of 6 to 8 weeks following radiation seems necessary to reduce or avoid false-positive results (78, 175). FDG-PET was sensitive (90%) and specific (83%) for persistent cancer after neoadjuvant chemotherapy in a 28-patient study (176). In a recent study of 41 patients treated definitively by radiation (with or without chemotherapy), FDG-PET findings were highly correlated with lymph node pathology, suggesting that neck dissection could be avoided in patients with negative FDG-PET scan (SUV <3.0; ref. 177). Moreover, in a separate study of patients receiving radiation with or without chemotherapy, a post-therapy metabolic rate below versus above the median was associated with a 5-year OS of 72% versus 35%, respectively (P = 0.0042; ref. 178).

**Lymphoma.** Although 30% to 70% of patients with advanced or aggressive lymphomas can be cured with first-line therapy, many still die of their disease. As detailed in Table 2, considerable evidence supports the significant association of post-therapy FDG-PET results with outcome. In particular, FDG uptake has been a significant early predictor of residual or recurrent disease and disease progression as well as PFS and OS (179–185). FDG-PET is especially useful in differentiating tumor from fibrosis within residual radiographic masses (186, 187). Such masses are present in half (non–Hodgkin’s lymphoma) to two-thirds (Hodgkin’s disease) of patients, of whom only 25% (non–Hodgkin’s lymphoma) to 30% (Hodgkin’s disease) eventually relapse. In a prospective study of 58 patients (43 with Hodgkin’s disease) with residual masses following treatment, FDG-PET (SUV <3) predicted recurrence (P = 0.004) and PFS (P < 0.00001; ref. 185). Disease progression was observed after 2 months in 16 of 19 versus 3 of 22 lymphoma patients with positive and negative FDG-PET results, respectively (P < 0.001; ref. 184). FDG-PET results obtained after the first cycle of chemotherapy in 30 non–Hodgkin’s lymphoma and Hodgkin’s disease patients predicted PFS at 18 months (P ≤ 0.001; ref. 188). A recent meta-analysis found that persistence of FDG-avid lesions after therapy predicted relapse, with up to 100% of patients with positive FDG-PET scans experiencing recurrence within 2 years (189). Conversely, absence of disease by FDG-PET scan is an indicator of a favorable prognosis. Indeed, the negative predictive value of FDG-PET was 96% in one study of 81 Hodgkin’s disease patients (106), and several studies have also reported a higher positive predictive value for FDG-PET versus conventional imaging modalities (106, 190–192). Integrating FDG-PET into the International Workshop response criteria also seems to increase the predictability of outcome (193, 194).

Prognosis is usually poor in nonresponding and relapsed patients regardless of further conventional treatment, but early, intensive treatment (e.g., high-dose chemotherapy with autologous stem cell transplantation) may be of benefit to appropriately selected patients. Timely identification of nonresponders (e.g., during induction therapy) is important for planning additional treatment of these patients. FDG-PET results following one to four cycles of therapy seems to predict outcome in patients with Hodgkin’s disease and aggressive non–Hodgkin’s lymphoma; however, whether altering therapy based on those observations will improve outcome has yet to be shown (180).

**Sarcoma.** The response to treatment in sarcoma is difficult to objectively measure and quantify anatomically as shown by the limited usefulness of the Response Evaluation Criteria in Solid Tumors in this setting (195, 196). Assessment of tumor dimensions in sites, such as bone, bowel, and peritoneal metastases, is problematic; in addition, tumor volume reductions that can be measured by standard criteria may occur slowly or not at all (e.g., due to persistence of necrotic or fibrotic tissue). Several studies now suggest that FDG-PET has significant potential for assessing response to treatment in sarcoma (197–204) as well as for detecting local relapse (200). In high-grade soft tissue sarcomas, chemotherapy remains controversial because response only approximates 40% (205). A meta-analysis of existing studies has not shown a definite benefit of FDG-PET imaging in the management of sarcomas (206); however, this is likely due to the inclusion of studies using poor methodologies as well as limited adherence to appropriate definitions of tumor response. FDG-PET has shown particular promise in monitoring sarcoma and gastrointestinal stromal tumor therapy with the targeted cytostatic agent imatinib (207, 208) and supported in part the demonstration of the agent’s efficacy in gastrointestinal stromal tumor (209–211). Compared with standard CT imaging, early therapy monitoring with FDG-PET was a better predictor of long-term outcome. For instance, FDG-PET imaging accurately predicted 1-year tumor response to imatinib in 85% (1-month scan) or 100% (3- or 6-month scan) of 20 gastrointestinal stromal tumor patients compared with only 57% (at 6 months) by CT (207). In a separate study of 21 gastrointestinal stromal tumor and soft tissue sarcoma patients, response based on FDG-PET data and as defined by the European Organization for Research and Treatment of Cancer criteria (ref. 212; see Table 3) obtained 8 days after imatinib treatment correlated with symptom control as well as longer PFS (ref. 208; see Fig. 5). Furthermore, in a recent study of only high-grade or large intermediate-grade soft tissue sarcomas, a 40% reduction in SUV_max following neoadjuvant chemotherapy was a statistically significant independent predictor of both DFS and OS (213).

**Other cancers.** Considerable evidence supports the utility of FDG-PET for monitoring response to therapy in breast cancer (214–222), and it has specifically been approved by CMS for...
reimbursement when a change in therapy is anticipated. Emerging data also suggest a role in cervix (142, 143, 223), gastric (224), colorectal (225–227), and testicular (136) cancers. In a small group of breast cancer patients responding to chemohormonotherapy, Wahl et al. reported a rapid and significant drop in the tumor FDG-PET SUV that was observable 8 days after initiating treatment (214). Numerous studies have subsequently confirmed the utility of FDG-PET for monitoring response to therapy in breast cancer, particularly in locally advanced or metastatic disease (215–222). In many reports, FDG-PET reliably differentiated responders from non-responders as early as after the first cycle of chemotherapy. FDG-PET is particularly useful for assessing response to neoadjuvant therapy when no change in tumor size occurs. FDG-PET may also be useful in monitoring early drug reactions, such as the flare response to antiestrogens (228, 229). In tamoxifen-treated postmenopausal women, increased FDG uptake 7 to 10 days after treatment was a good predictor of objective clinical response based on 3 to 24 months of follow-up even in the absence of a clinical flare response (229).

In several settings (e.g., colorectal and cervical), the utility of FDG-PET for initial staging and restaging disease, particularly when tumor volume is unchanged or changes slowly, also has advantages for assessing treatment response. For example, as noted in Table 2, FDG-PET is significantly predictive of response to therapy in cervical cancer. This is because FDG-PET is relatively accurate for assessing the extent of disease, particularly in lymph nodes that do not change in size following therapy. The extent of lymph node metastases determined by FDG-PET predicted 3-year cause-specific survival in 47 treated stage IIIb patients (223). Compared with a 73% survival at 45 months in patients negative for lymph node FDG uptake, those with increased FDG signal from pelvic, plus para-aortic, or plus para-aortic and supraclavicular nodes had reduced survival rates of 58%, 29%, and 0%, respectively ($P = 0.0005$). In a retrospective study of 76 newly diagnosed cervical cancer patients, OS was 30% and 70%, respectively, for those with any or no post-treatment FDG uptake in the cervix and lymph nodes (142). None of the patients who developed new sites of FDG uptake were alive at 2 years. As confirmed in a later report from the same investigators of an expanded population of 152 cervical cancer patients (143), post-treatment FDG-PET was the most significant prognostic factor for death from cervical cancer.

Accumulating evidence supports the utility of FDG-PET for assessing therapy response in colorectal cancer as a predictor of long-term outcomes. In rectal cancer, neoadjuvant therapy can enhance the length as well as the quality of life, the latter due to improved pelvic control and sphincter preservation. Emerging data indicate that, compared with anatomic imaging modalities, FDG-PET can better differentiate scar tissue from locally recurrent rectal cancer and thereby improve response assessment. For example, a recent study of 15 rectal cancer patients found a significantly larger mean change in SUV$_{\text{max}}$ (69% versus 37%) for patients remaining free from recurrence following presurgical chemoradiation, with a larger change in SUV$_{\text{max}}$ ($\geq 62.5$) correlated with increased disease-specific and recurrence-free survival (226). FDG-PET has high sensitivity for colorectal cancer recurrence (230) and has particular utility in assessing the response in colorectal patients with hepatic metastases. For example, response to 5-FU with or without IFN was associated with lower FDG-PET SUVs at 4 to 5 weeks or with lower tumor/liver ratios at 1 to 2 or 4 to 5 weeks (227). Similarly, FDG-PET identified patients responsive to combination therapy with 5-FU and folic acid (231) or 5-FU, folic acid, and oxaliplatin (the FOLFOX regimen; ref. 232). Finally, several studies have shown the utility of FDG-PET for monitoring response to local ablative therapy for colorectal liver metastases (233–236).

The Developmental Path for Validation of FDG-PET as a Surrogate Marker for Clinical Benefit and Its Value in Oncologic Drug Development

Taken together, studies to date establish a role for FDG-PET in assessing response to standard therapies and predicting outcome. These data suggest that FDG-PET has potential to be validated as a surrogate end point for clinical benefit. Once validated with approved therapies, FDG-PET could be employed as a trial end point both in phase III accelerated approval trials and to support go/no go decisions in phase II clinical trials. As such, FDG-PET has the potential to accelerate the drug development process by allowing dosing adjustments or early identification of responders. The paragraphs below present seven case studies, which highlight the outstanding issues and drug development opportunities of employing FDG-PET as a surrogate marker of clinical benefit. Based on existing data, European Organization for Research and Treatment of Cancer has published recommendations regarding the use of FDG-PET for disease assessment (see Table 3; ref. 212). It is anticipated that further insight into the appropriate target organ-specific cutoffs and application of FDG-PET will be defined using receiver operator characteristic analyses in specific cancer types. Once available, these data will guide the design of definitive prospective validation studies of FDG-PET for clinical benefit. As an example, Fig. 3 shows such a prospective validation study of FDG-PET with standard approved chemotherapy in NSCLC. Once validated, FDG-PET could then be incorporated into studies of new therapeutics to accelerate their development and ultimately facilitate progress in the management of cancer patients.

Non–small cell lung cancer. Because treatment failure is closely followed by death in NSCLC, it has been possible to

### Table 3. Proposed European Organization for Research and Treatment of Cancer criteria for assessment of response by FDG-PET

<table>
<thead>
<tr>
<th>Category</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progressive metabolic disease</td>
<td>Increase of SUV &gt; 25%</td>
</tr>
<tr>
<td>Stable metabolic disease</td>
<td>Increase of SUV (25% or decrease (15%)</td>
</tr>
<tr>
<td>Partial metabolic response</td>
<td>Reduction of a minimum of 15-25% of SUV after one treatment cycle; 125% after more than one treatment cycle</td>
</tr>
<tr>
<td>Complete metabolic response</td>
<td>Complete resolution of FDG uptake</td>
</tr>
</tbody>
</table>

Non–small cell lung cancer. Because treatment failure is closely followed by death in NSCLC, it has been possible to...
correlate the FDG-PET response after a single cycle of chemotherapy with patient outcome. As shown in Fig. 4, Kaplan-Meier survival curves indicate that response as assessed by a single, early FDG-PET scan is highly predictive of survival (161). Compared with CT imaging, response assessment by FDG-PET can be conducted much sooner and serial scans (i.e., to assess residual anatomic masses) are not usually required. Moreover, FDG-PET was superior to CT for predicting survival in a recent study of 73 patients (163) and in the interim analysis of an ongoing multicenter trial (reviewed in ref. 50). In addition, in the erlotinib (Tarceva) randomized trial, the survival benefit of 2 months was not accounted for solely by the responders identified by anatomic imaging, showing that patients with stable disease contributed to the outcome. This suggests that molecular imaging like FDG-PET can refine the classic response criteria, particularly for cytostatic agents. These encouraging data suggest that the FDG-PET response can be validated retrospectively against survival and other end points to determine appropriate SUV cutoffs for defining response. One approach is a quantitative comparison of FDG-PET response with end points used to show clinical benefit in the approval trials of NSCLC treatments. Approved therapies for first-line treatment of NSCLC include four combination cisplatin therapies and single-agent vinorelbine. The approval bases were superior survival (3), noninferior survival (1), and a superior TTP response rate with a survival trend (1). Benefit ranged from an 18% to 36% increase in OS with absolute survival increases from 2 to 2.6 months versus the active comparator arms. The observed 66% increase in survival benefit (3.4 months absolute value) observed in the FDG-PET study by Weber et al. (161) is comparable with the response seen for these same drugs in their pivotal approval trials. This provides an opportunity to validate the FDG-PET end point and the proposed SUV cutoff values retrospectively (e.g., using receiver operator characteristic analyses). These data, in aggregate, indicate the likelihood that an efficient validation study as shown in Fig. 3 will be successful in lung cancer; similar validation trial designs can be developed for other target organs. FDG-PET could then be used to enhance development of new oncologic drugs.

Lymphoma. In lymphoma, complete clinical responses are seen more frequently than in solid tumors; they correlate with, and are accepted surrogates of, survival. However, shorter end points are needed, particularly in phase II trials conducted in relapsed patients and those with refractory disease (193). A key issue in lymphoma is the post-treatment characterization of residual masses to discriminate cancerous from necrotic or fibrotic tissue, and several studies have now established the high predictive value of FDG-PET compared with anatomic imaging (186). These data suggest that the definition of a clinical response by anatomic imaging should be refined with a confirmatory FDG-PET scan to detect residual disease in those patients with normal-sized nodes or to rule out disease in those with enlarged nodes. Moreover, treatment-induced changes in FDG uptake have been observed within 1 to 3 days; SUVs after one cycle of chemotherapy correlate well with established outcome measures (188). Several investigations have found that lack of response using FDG-PET criteria (e.g., persistent FDG uptake or <25% decrease) correlated with PFS and OS in both Hodgkin's disease and non–Hodgkin's lymphoma (180–185, 188, 190, 237). FDG-PET has a high negative predictive value (96% in one 81-patient study; ref. 106) and a higher positive predictive value than conventional imaging modalities (106, 190–192). Thus, the data to date have established FDG-PET as an acceptable surrogate end point for clinical benefit in lymphoma. Opportunities for its future utility include refinement of classification of clinical responses currently defined by

\[ \text{FDG-PET & CT image analysis at diagnosis (SUV calculation, standard imaging measurements).} \]
\[ \text{Sample collection: Blood/serum/sputum (surrogate PBL or bronchoalveolar cellular or proteomic/molecular target response), FNA, biopsy or surgical resection sample (standard histopathology or immunocytochemistry or gene expression array analysis); Standard clinical data (history, performance status, symptomology); demographic information (e.g., age, sex, race).} \]
\[ \text{1st line therapy (platinum-, taxane-, microtubule inhibitor-based) or novel cytotoxic agent therapy or combination + 1st line therapy.} \]
\[ \text{Multiple cycles of treatment as supported by preclinical, early clinical, and MTD rationale with regular FDG-PET & CT image analysis to determine metabolic response.} \]
\[ \text{Stratification by responders vs. non-responders (FDG-PET SUV decrease 20-35%); Determination of standard CT-based RR} \]
\[ \text{Post-treatment: Randomize FDG-PET/standard CT responders to continued novel therapy or standard + novel therapy vs. placebo.} \]

**Trial endpoints**
- Biomarker analyses (e.g., molecular target in surrogate tissue/primary tumor, standard histopathology grading, proteomic, gene expression microarray analysis, etc.)
- Response measure (disease-free survival, progression-free survival, duration of response, objective response rate, etc.)
- Safety/tolerability of treatment
- Survival
- Quality of Life

**Fig. 3.** Validation of FDG-PET metabolic response as an approvable surrogate end point in a randomized discontinuation study of a novel cytostatic for the treatment of NSCLC. The recent study by Weber et al. (161) established that a prospectively defined FDG-PET response predicts response to and survival following, cytotoxic chemotherapy in NSCLC (see Fig. 4). Based on these data, prospective trial designs are possible to allow exploration of the further utility and evaluation of FDG-PET with existing and novel therapies.
anatomic imaging and its use in measuring response in phase II trials of novel lymphoma therapies.

Breast. As detailed in the preceding sections, FDG-PET is approved by CMS to monitor response to breast cancer therapy when a change in therapy is anticipated. Several studies support its accuracy for assessing response to cytotoxics as well as antiestrogens. One breast cancer setting for further validation of FDG-PET in assessing new oncologic agents would be in evaluating neoadjuvant therapy in patients with early-stage or locally advanced disease scheduled for surgery. Such a trial would provide opportunity for validation of FDG-PET against the histologic end point of extent of residual disease in the breast tissue at definitive surgery. Once validated, FDG-PET could be employed as an end point in phase II trials, assessing response in this neoadjuvant setting as a predictor of systemic benefit, in eradicating micrometastatic disease. Data are emerging that a pathologic complete response in the breast following neoadjuvant chemotherapy predicts systemic benefit and a low rate of systemic recurrent disease (238). In addition, FDG-PET also has application as a surrogate end point in several other clinical breast cancer settings. In patients with stage IIIa or IIIb disease, FDG-PET could be used as an early predictor of drug efficacy for both local and systemic diseases. In such patients, 6-month preoperative anthracycline- and taxane-based chemotherapy is standard. An indication by FDG-PET of a lack of antitumor activity could signal discontinuation of toxic, ineffective therapy and initiation of alternative, non-cross-resistant therapies. Another setting where FDG-PET has application is as an early indicator of drug efficacy in estrogen receptor–positive metastatic disease. Aromatase inhibitors, such as anastrozole (Arimidex), letrozole (Femara), or exemestane (Aromasin), have shown efficacy in such patients. FDG-PET
could be used as an early end point in trials assessing novel agents tested in combination with, or as a comparator to, these standard therapies. Finally, FDG-PET has high promise in aggressive breast cancer, often found in younger women. This suggests the need for correlative studies of FDG-PET with prognostic and predictive factors (e.g., premenopausal, estrogen receptor–negative, HER-2-positive, and BRCA1 status and proliferative rate).

Prostate. To date, FDG-PET studies in prostate cancer have typically included heterogeneous populations, with mixed findings (e.g., refs. 239–241). Emerging data have begun to define the disease settings in which FDG-PET has value in detecting recurrence and its potential to serve as a short-term indicator of response. For localized disease and recurrence, MRI offers superior resolution than FDG-PET (242). Lymphotrophic superparamagnetic nanoparticle-facilitated MRI imaging and existing nomograms will likely prove better than FDG-PET for nodal disease (243). In patients with metastatic disease who are to receive androgen deprivation, prostate-specific antigen is typically a better (and much less costly) measure of disease; however, FDG-PET could potentially be used instead of bone scintigraphy to identify active, nonproliferating or silent lesions versus those that continue to proliferate and which may be appropriate for consolidation therapy. FDG-PET also has utility in patients with rising prostate-specific antigen after radiation or radical prostatectomy; prostate-specific antigen doubling time is of prognostic value (244–246), and FDG-PET could provide complementary information, including the identification of metastatic disease sites. This contrasts with MRI, which may be better for local disease. FDG-PET is also of value in patients with hormone-refractory prostate cancer, in whom bone metastases are the primary cause of morbidity and mortality. Their development usually signifies the transition to the lethal phase of the illness. Recent data suggest that the modality has utility for monitoring these metastatic bone lesions (247, 248).

Patients with hormone-refractory prostate cancer have limited expected survival and few treatment options, and the continued identification of active agents is essential. The recent approval of taxotere, in combination with prednisone, based on demonstration of a 2.5-month survival benefit in phase III testing has stimulated testing of three- and four-drug regimens (249) and combinations of taxotere with agents that have shown synergy in preclinical studies. The latter include calcitriol (250), oblimersen sodium (Gensense, an antisense oligonucleotide that binds to BCL-2 mRNA; ref. 251), and the angiogenesis inhibitors thalidomide (252) and bevacizumab (Avastin, an antibody to vascular endothelial growth factor); increased vascular endothelial growth factor has prognostic significance in this patient group (253). Similarly, the plateau-derived growth factor inhibitor imatinib is being studied based on immunohistochemical and in vivo studies showing high expression in bone metastases (254). Although it has limited activity as a single agent (255), responses were observed when docetaxel was added (256). Synergy has also been observed in preclinical models with the ansamycin 17-allylamino geldanamycin (257, 258), which results in degradation of heat shock protein 90 client proteins, including HER-2 and the androgen receptor. The combination of 17-allylamino geldanamycin and docetaxel is therefore undergoing clinical testing; combinations with signal transduction inhibitors (trastuzumab, bevacizumab, etc.) also have promise. Given limited patient resources and the cost of conducting definitive studies, rigorous criteria are urgently needed to determine whether, and to what degree, new single and combination therapies are more efficacious than taxotere alone to make the decision to proceed to phase III testing objectively. One approach is to estimate the benefit of a novel therapy relative to a historical group using nomogram-based predictions of survival (259). Another avenue is to employ surrogate end points to facilitate assessment of novel agents or new taxotere combinations in small, comparative phase II trials. Assessment of metastatic bone lesions, even in asymptomatic patients, would represent a key measure of drug efficacy. The promise of FDG-PET as an outcome measure in this setting is already being evaluated; a 25% increase in the average SUV_max for FDG 4 weeks after antimicrotubule chemotherapy accurately identified the clinical status (based on prostate-specific antigen, bone scintigraphy, and soft tissue imaging) in 91% (20 of 22) of patients (260). Ongoing effort is aimed at refining SUV cutoffs. It is anticipated that, with further prospective validation against survival outcomes, FDG-PET could be a valuable early end point to facilitate efficacy assessments in this metastatic disease setting and, combined with prostate-specific antigen–based end points, provide a reasonably likely measure of clinical benefit.

Sarcoma. Sarcoma is a relatively rare but deadly disease. Osteogenic and Ewing's sarcomas affect mostly children, although both chondrosarcoma and soft tissue sarcomas predominantly occur in adults (261). Surgical resection is the mainstay of treatment, and for many histologies, presurgical chemotherapy and/or radiation therapy are also administered. In the United States, most children with cancer are treated definitively on clinical trials (262). This is due to the recognized need among pediatric oncologists for cooperative clinical research as well as the likelihood that childhood cancers will respond to treatment. Sufficiently powered clinical trials using survival as an end point are extremely challenging (205, 262, 263), and there is a clear need to assess the early clinical signals of the efficacy of candidate drugs. End points that might serve as a surrogate to survival would facilitate identification and early testing of the most promising novel agents or therapies, thereby conserving the precious resources and time required for large clinical trials for only these candidates. The histologic response to neoadjuvant treatment correlates with survival in osteosarcoma, Ewing's sarcoma, and soft tissue sarcomas (264). As histologic response can only be assessed after surgical excision of a tumor, a noninvasive means is needed to assess ongoing treatment response and to use as a surrogate for survival. Data from single institution studies suggest that the FDG-PET signal correlates with tumor response to neoadjuvant treatment and is predictive of survival in both bone and soft tissue sarcomas (173, 197, 213, 265). As shown in the Kaplan-Meier survival curves in Fig. 5, FDG-PET responders experienced a significantly improved PFS; 1-year PFS was 92% in responders compared with only 12% in nonresponders (208). The need for multicenter trials to confirm these findings is paramount to validate FDG-PET as a surrogate measure for survival. This would serve to both enable and accelerate needed oncologic drug development for this serious unmet medical need and undoubtedly would facilitate progress in the care of patients with sarcomas.

Colorectal. At a recent Oncologic Drug Advisory Committee meeting (266), DFS was accepted as a surrogate end point in the adjuvant setting. In metastatic disease, it is not yet clear that PFS...
or TTP is an acceptable surrogate of clinical benefit to support drug approvals. In five recent approvals in colon cancer, disease progression has been assessed using standard imaging techniques (e.g., Response Evaluation Criteria in Solid Tumors). Based on its improved sensitivity and specificity for detecting and characterizing recurrent and metastatic disease, FDG-PET could improve the sensitivity of time-to-recurrence measures in advanced colon cancer patients. FDG-PET data may be available at shorter intervals than conventional imaging, allowing more immediate assessments of drug efficacy. FDG-PET could also refine the determination of eligibility for hepatic resection based on absence of extrahepatic metastatic disease (93, 94). FDG-PET should be incorporated as an end point into clinical trials to assess its value for predicting outcome as assessed by standard imaging and survival. Finally, FDG-PET could be considered for assessments of rectal cancer. Recurrence or 3-year local control

Fig. 5. FDG-PET response to imatinib predicts PFS in advanced soft tissue sarcomas. Kaplan-Meier survival curves for all patients (A) and FDG-PET responders versus nonresponders (B). Reprinted with permission from ref. 208.
correlates with clinical benefit, and emerging data suggest the utility of FDG-PET for assessing response to therapy in rectal cancer (e.g., refs. 145, 225, 226).

**Ovary.** Emerging data suggest that FDG-PET has utility for detecting and measuring recurrent or residual disease, particularly in women with elevated or rising CA-125. CA-125 has already been suggested as an important surrogate end point for drug development studies in ovarian cancer (267). Just as the initial CMS approval of FDG-PET in colon cancer targeted patients with an unexplained increase in carcinoembryonic antigen, a rising CA-125 could be used to select patients for follow-up imaging with FDG-PET. This approach could lead to an improved characterization of disease in women with a rising or elevated CA-125 and could be valuable as a confirmatory end point in trials. As an example, a trial could be designed using the CA-125 response, with patients experiencing a increase in CA-125 screened with FDG-PET to detect recurrence. The SUV cutoffs proposed by the European Organization for Research and Treatment of Cancer could be used initially to define the FDG-PET-based response in this setting.

### Conclusions and Recommendations

FDG-PET has been used to image cancers for >25 years, with approval for Medicare reimbursement beginning in 1998. Accumulating evidence supports the value of the methodology as an essential tool for guiding patient care. As an oncology imaging modality, FDG-PET is approved for Medicare reimbursement in 10 clinical settings (see Table 1); in all other cancers, coverage applies only to FDG-PET scans conducted in certain prospective trials or patient registries. In contrast to conventional imaging methodologies, FDG-PET provides information about the rate of glycolytic metabolism, rather than the anatomic structure, of the cancer. The approaches are thus complementary, and in practice, FDG-PET is used in most settings to stage and restage pathologically proven cancers that have been identified by radiography, CT, or MRI. FDG-PET has particular utility for identifying or characterizing FDG-avid lesions that cannot be readily visualized or discriminated from normal tissue with conventional methodologies. For example, FDG-PET can locate unknown primary head and neck or non-181-avid thyroid tumors and can detect and characterize nodal disease or distant metastases (e.g., in melanoma, esophageal, and colorectal cancers). In many cancers, FDG-PET also has prognostic utility. As indicated by CMS approval of FDG-PET in only defined settings, FDG-PET data are essential for clarifying certain clinical decisions, but it is not necessary to perform FDG-PET scans in all cases. Ongoing and additional studies will more clearly define the settings and circumstances in which FDG-PET scans are most appropriate.

FDG-PET is more accurate than conventional imaging for restaging many cancers after treatment and, for this reason, is well suited to application in therapeutic monitoring. Although only approved for this use in breast cancer, several studies now support the utility of FDG-PET as an early indicator of response to chemotherapy or radiotherapy in a broad range of cancers. In settings such as NSCLC and esophageal cancer, the FDG-PET signal decline in response to therapeutic intervention well before changes in tumor size that can be visualized using conventional imaging are anticipated. Because it is critical to planning surgical resection or further therapeutic intervention, FDG-PET has the potential to greatly impact the management of cancer patients by providing a more accurate and timely assessment of response to neoadjuvant or definitive treatment. In addition, FDG-PET can also be applied to facilitate greater efficiency in evaluating drug efficacy in clinical trials. A key step in this process is the validation of FDG-PET against accepted measures of response, including conventional imaging assessments as well as survival outcomes. Accumulating data already support the validation of FDG-PET for response assessment in several settings, and prospective validation trials in other settings could be readily achieved by incorporating FDG-PET end points into phase II or III trials. Further studies to explore and define the application of FDG-PET as an early surrogate for clinical benefit are warranted. The following summarizes the data supporting the potential of FDG-PET to facilitate oncologic drug development and provides specific recommendations to validate and implement the approach.

- Cancer cells use glycolysis for energy production. Indeed, an elevation in glucose transporters, glucose consumption, or glycolytic enzymes has been correlated with tumor grade, prognosis, or treatment response in some cancers. Although the regulation of genes encoding glucose transporters and hexokinases has not been fully elucidated, oncogenic signal transduction pathways can directly stimulate transcription of these genes and trigger glycolysis. In particular, the protein kinase Akt seems to be one of the key mediators of the establishment and maintenance of glycolysis in cancer cells.

- FDG-PET exploits the reliance of tumor cells on glucose to image cancers. FDG is a glucose analogue that becomes trapped intracellularly on injection, uptake, and phosphorylation. As the $^{18}$F in accumulated FDG decays, it is detected with a PET scanner. As is supported by a considerable body of literature, FDG-PET is a specific, sensitive, and reproducible imaging modality for cancer and has been widely applied in oncologic settings.

- FDG-PET data can be assessed visually or analyzed semiquantitatively or quantitatively. Quantitative derivation of the metabolic rate is the more robust approach but requires dynamic scanning as well as an arterial input function. However, neither dynamic scanning nor arterial input are required to define the SUV, a semiquantitative index of tumor uptake normalized to injected dose and subject’s body weight (or lean body mass or body surface area). The SUV is strongly associated with the quantitatively derived metabolic rate and is an accepted and widely used measure of FDG uptake. The SUV offers a suitable tradeoff between precision and simplicity and is of particular value in assessing differences among scans where cancer patients serve as their own control. However, the SUV, as an approximation, can lead to bias in the estimates of glucose metabolism unless there are standardized imaging protocols and approaches to image analysis.

- There is a need to standardize the methodology for FDG-PET data collection and analysis, and National Cancer Institute plans a series of workshops to address this. One key issue is the definition of SUV cutoffs (absolute levels and expected percentage changes) that would discriminate nonneoplastic processes from cancer or tumor progression from stable disease from response. SUV cutoff values have been proposed in defined disease settings. Further research in this area would be facilitated by retrospective
In several clinical settings (e.g., NSCLC, esophageal cancer, and lymphoma), FDG-PET can provide an early measure of response to treatment with approved therapies. Understanding the tumor response is an essential consideration in patient management (e.g., discontinuing ineffective therapies or planning additional surgical or therapeutic intervention); thus, FDG-PET may significantly affect patient outcome. The FDG-PET response was well correlated with conventional measures of disease progression and survival, and FDG-PET was superior to anatomically based imaging modalities for assessing response and predicting outcome in some studies. These emerging data are sufficient to validate FDG-PET as a surrogate end point of clinical benefit in some cases; in others, initial or additional retrospective studies (e.g., with receiver operator characteristic analyses) are needed to define appropriate SUV cutoffs and the anticipated magnitude of response. These data can then be used to design prospective studies that will provide definitive validation of FDG-PET end points.

Once validated, the application of FDG-PET as a surrogate for clinical benefit has the potential to facilitate drug development. For example, the modality may shorten the duration of phase II studies. Phase III trials with a FDG-PET end point could serve as the basis for accelerated approval, with full approval contingent on evidence of a survival benefit with longer-term follow-up. The seven case studies (NSCLC, lymphoma, breast, prostate, sarcoma, colon, and ovary) presented in this article highlight the opportunities for a much expanded use of FDG-PET in drug development and to support drug approvals. These include clinical settings where current measures of treatment efficacy and disease progression are inadequate, in which FDG-PET may provide a superior and earlier assessment of drug efficacy.

Acknowledgments

The authors gratefully acknowledge Dr. Greg Baxter and Barbara Gress of CCS Associates.

References

22. Zhong H, Chiles K, Feldser D, et al. Modulation of...


gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity to gefitinib or erlotinib. Proc Natl Acad Sci U S A 2004;101:13306–11.


175. Pao W, Miller V, Zakowski M, et al. EGFR receptor

Clin Cancer Res 2005;11(8) April 15, 2005 2806

Downloaded from clinicanres.aacrjournals.org on April 16, 2017, 2005 American Association for Cancer Research.


60. Larson SM, Morris M, Gunther I, et al. Tumor www.aacrjournals.org

2807 Clinical Cancer Res 2005;11(8) April 15, 2005

Downloaded from clinicanres.aacrjournals.org on April 16, 2017. © 2005 American Association for Cancer Research.


Progress and Promise of FDG-PET Imaging for Cancer Patient Management and Oncologic Drug Development

Gary J. Kelloff, John M. Hoffman, Bruce Johnson, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/11/8/2785

Cited articles
This article cites 256 articles, 102 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/11/8/2785.full.html#ref-list-1

Citing articles
This article has been cited by 63 HighWire-hosted articles. Access the articles at:
/content/11/8/2785.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.