

# Recommendations on the Use of $^{18}\text{F}$ -FDG PET in Oncology

James W. Fletcher<sup>1</sup>, Benjamin Djulbegovic<sup>2</sup>, Heloisa P. Soares<sup>2</sup>, Barry A. Siegel<sup>3</sup>, Val J. Lowe<sup>4</sup>, Gary H. Lyman<sup>5</sup>, R. Edward Coleman<sup>5</sup>, Richard Wahl<sup>6</sup>, John Christopher Paschold<sup>7</sup>, Norbert Avril<sup>8</sup>, Lawrence H. Einhorn<sup>1</sup>, W. Warren Suh<sup>9</sup>, David Samson<sup>10</sup>, Dominique Delbeke<sup>11</sup>, Mark Gorman<sup>12</sup>, and Anthony F. Shields<sup>13</sup>

<sup>1</sup>Indiana University School of Medicine, Indianapolis, Indiana; <sup>2</sup>H. Lee Moffitt Cancer Center at University of South Florida, Tampa, Florida; <sup>3</sup>Washington University School of Medicine, St. Louis, Missouri; <sup>4</sup>Mayo Clinic, Rochester, Minnesota; <sup>5</sup>Duke University Medical Center, Durham, North Carolina; <sup>6</sup>Johns Hopkins Medical Center, Baltimore, Maryland; <sup>7</sup>US Oncology, Newport News, Virginia; <sup>8</sup>Queen Mary's School of Medicine and Dentistry, London, United Kingdom; <sup>9</sup>Brigham and Women's Hospital, Boston, Massachusetts; <sup>10</sup>Blue Cross and Blue Shield Association, Washington, DC; <sup>11</sup>Vanderbilt University Medical Center, Nashville, Tennessee; <sup>12</sup>National Coalition for Cancer Survivorship, Washington, DC; and <sup>13</sup>Karmanos Cancer Institute, Detroit, Michigan

The rationale was to develop recommendations on the use of  $^{18}\text{F}$ -FDG PET in breast, colorectal, esophageal, head and neck, lung, pancreatic, and thyroid cancer; lymphoma, melanoma, and sarcoma; and unknown primary tumor. Outcomes of interest included the use of  $^{18}\text{F}$ -FDG PET for diagnosing, staging, and detecting the recurrence or progression of cancer. **Methods:** A search was performed to identify all published randomized controlled trials and systematic reviews in the literature. An additional search was performed to identify relevant unpublished systematic reviews. These publications comprised both retrospective and prospective studies of varied methodologic quality. The anticipated consequences of false-positive and false-negative tests when evaluating clinical usefulness, and the impact of  $^{18}\text{F}$ -FDG PET on the management of cancer patients, were also reviewed. **Results and Conclusion:**  $^{18}\text{F}$ -FDG PET should be used as an imaging tool additional to conventional radiologic methods such as CT or MRI; any positive finding that could lead to a clinically significant change in patient management should be confirmed by subsequent histopathologic examination because of the risk of false-positive results.  $^{18}\text{F}$ -FDG PET should be used in the appropriate clinical setting for the diagnosis of head and neck, lung, or pancreatic cancer and for unknown primary tumor. PET is also indicated for staging of breast, colon, esophageal, head and neck, and lung cancer and of lymphoma and melanoma. In addition,  $^{18}\text{F}$ -FDG PET should be used to detect recurrence of breast, colorectal, head and neck, or thyroid cancer and of lymphoma.

**Key Words:** oncology; PET;  $^{18}\text{F}$ -FDG PET

**J Nucl Med 2008; 49:480–508**

DOI: 10.2967/jnumed.107.047787

**P**ET is an imaging technique that provides unique information about the molecular and metabolic changes associated with disease. The technology has existed for more than 30 years but has been used clinically for only the last 10–15 years. In this period, dramatic improvements in technology, the routine availability of medical cyclotrons (to produce the necessary short-lived positron emitters), and favorable reimbursement decisions in the late 1990s have led to a tremendous increase in the use of this technology. The major area of clinical application is currently in oncology, with some application in cardiology and neurology.

PET requires the use of molecules (radiopharmaceuticals) that are labeled with radioactive nuclides. The amounts of radiolabeled material administered are extremely small ( $10^{-6}$ – $10^{-9}$  g) and have essentially no pharmacologic effect. In this regard, PET has the unique ability to assess molecular alterations associated with disease without perturbing or altering the fundamental underlying molecular and biochemical processes. Although the number of molecular probes that can be radiolabeled with positron emitters is extremely large, and clinical investigational uses number in the thousands, clinical practice has been limited principally to the use of a glucose analog labeled with the positron emitter  $^{18}\text{F}$ -FDG.  $^{18}\text{F}$ -FDG was first synthesized in 1978 (1) and has become the most commonly used radiopharmaceutical for PET studies of cancer and also for the study of normal functions and diseases of the brain and heart. In March 2000, the Food and Drug Administration approved the use of  $^{18}\text{F}$ -FDG to assist in the evaluation of malignancy in patients with known or suspected abnormalities found by other testing methods or in patients with an existing diagnosis of cancer.

The fact that cancer cells exhibit an increased rate of glycolysis has been known since the 1920s (2), and  $^{18}\text{F}$ -FDG PET is able to assess a fundamental alteration in the cellular metabolism of glucose that is common to all

Received Oct. 1, 2007; revision accepted Nov. 20, 2007.

For correspondence or reprints contact: James W. Fletcher, MD, Department of Radiology, Indiana/Purdue University, Indiana University School of Medicine, University Hospital, Room 0655, 550 N. University Blvd. Indianapolis, IN 46202-5253.

E-mail: jwfletch@iupui.edu

COPYRIGHT © 2008 by the Society of Nuclear Medicine, Inc.

neoplasms. Increased cellular glucose uptake is one of the key alterations associated with the high glycolytic rate of cancer cells.

## HISTORY

The first medical application of positron emitters was reported more than 50 years ago in 1951 by Sweet at Massachusetts General Hospital (3). This application involved a simple probe that used coincidence detectors to localize tumors in the brain. The first published PET images were acquired using a ring tomograph with the filtered backprojection algorithm and included images of oxygen metabolism with  $^{15}\text{O}$ -oxygen and glucose metabolism with  $^{11}\text{C}$ -glucose, as well as  $^{18}\text{F}$ -fluoride bone images (4,5). This publication occurred in 1976, almost 25 years after Sweet's work at Massachusetts General Hospital. Significant subsequent advances in PET technology were associated with the identification of bismuth-germanium-oxide as a scintillator material in 1977 (6) and the successful synthesis of  $^{18}\text{F}$ -FDG by Ido et al. at Brookhaven in 1978 (1). The first  $^{18}\text{F}$ -FDG scans were obtained at the University of Pennsylvania in 1979 by Phelps et al. using  $^{18}\text{F}$ -FDG that was synthesized at Brookhaven National Laboratory in Long Island (7–9). The most recent technical innovation, which has been available for only the last few years, is the integration of PET and CT systems. These dual-modality systems offer an advantage over dedicated PET in that they can concurrently provide both metabolic and structural or anatomic images that are automatically fused and overcome some limitations of dedicated PET.

Reimbursement for PET procedures was not available through much of the 1990s, and adoption of the technology was slow. In 1995, the Food and Drug Administration approved  $^{18}\text{F}$ -FDG for brain imaging in patients with epilepsy. This approval paved the way for Health Care Financing Administration reimbursement of PET in January 1998 for lung cancer and cardiovascular disease in Medicare beneficiaries. This coverage was expanded by the Health Care Financing Administration in 1999 to include restricted indications for colorectal cancer, melanoma, and lymphoma. In the following year, the Food and Drug Administration gave broad approval for  $^{18}\text{F}$ -FDG in all cancers and cardiovascular disease. Near the end of 2000, the Health Care Financing Administration expanded coverage for broad use of  $^{18}\text{F}$ -FDG PET in lung, colorectal, head and neck, and esophageal cancers as well as lymphoma and melanoma. Since that time, indications have been added for breast cancer and thyroid cancer. In February 2006, the Centers for Medicare and Medicaid Services (the new agency name for the Health Care Financing Administration) announced that it would provide coverage for use of  $^{18}\text{F}$ -FDG PET in essentially all other cancers in accordance with its "coverage with evidence development" program. For Medicare beneficiaries undergoing PET as part of this program, referring physicians and PET facilities will be required to provide certain data to the

National Oncologic PET Registry to allow for assessment of the impact of PET on intended patient management.

## GENERAL LIMITATIONS OF DEDICATED $^{18}\text{F}$ -FDG PET

There are inherent limitations of  $^{18}\text{F}$ -FDG PET that can result in false-negative and false-positive findings. False-positive findings are most commonly associated with uptake of  $^{18}\text{F}$ -FDG in infectious or inflammatory tissue (10).  $^{18}\text{F}$ -FDG has been reported to accumulate in various inflammatory processes (11–13). Infection imaging with  $^{18}\text{F}$ -FDG PET relies on the fact that granulocytes and mononuclear cells use glucose as an energy source during and only during their metabolic burst (14,15), which takes place when activated by local triggers. It is therefore not surprising that  $^{18}\text{F}$ -FDG accumulates in many types of inflammatory tissue. For example,  $^{18}\text{F}$ -FDG uptake can be seen in tissue after radiation therapy. Inflammatory changes after radiation therapy can be protracted and a potential source of false-positive findings if the history, timing, and volume of tissue irradiated are not considered at the time of interpretation.  $^{18}\text{F}$ -FDG uptake can vary widely in normal tissue, and regions of discrete uptake in areas such as the ureters, bowel, lymphatic tissue, thymus, brown fat, and muscle—so called normal variants—can be interpreted in error as abnormal or can confound the correct interpretation of the findings. Mildly to moderately increased  $^{18}\text{F}$ -FDG uptake can also be seen in a variety of benign processes, many of which represent inflammatory or hyperplastic conditions (e.g., villous adenomas, thyroid adenomas, Graves disease, adrenal adenoma, Paget's disease, and fibrous dysplasia), and familiarity with the behavior of these and other conditions is important in diminishing false-positive results.

Weaknesses of  $^{18}\text{F}$ -FDG PET for cancer imaging include its limited reconstructed spatial resolution of 4–10 mm in available commercial systems. Negative scan findings cannot exclude the presence of a small tumor or microscopic tissue involvement, and precise anatomic localization of the signal can be difficult in certain anatomic regions (e.g., the head and neck). Tumors with a low metabolic rate (e.g., bronchoalveolar carcinoma and mucinous adenocarcinoma) may show minimal uptake of  $^{18}\text{F}$ -FDG, and certain tumors are known to have poor avidity for  $^{18}\text{F}$ -FDG (prostate carcinoma and hepatocellular cancer).  $^{18}\text{F}$ -FDG PET is also generally considered to not be useful in the assessment of possible cerebral metastases from known primary neoplasms. High levels of  $^{18}\text{F}$ -FDG are normally present in the cerebral cortex and substantially limit the utility of  $^{18}\text{F}$ -FDG PET in this application. For this reason, most clinical examinations are of the patient's torso and include the area from the base of the brain to the mid thigh.

## RATIONALE FOR THE RECOMMENDATIONS

The adoption of PET has been variable, but despite limitations in the published literature,  $^{18}\text{F}$ -FDG PET is

rapidly becoming an integral part of oncology practice in the United States, Europe, and other countries.

For these reasons, a multidisciplinary expert panel of oncologists, radiologists, and nuclear physicians with expertise in PET/CT convened to develop recommendations on the use of  $^{18}\text{F}$ -FDG PET in oncology practice and to determine the suitability of  $^{18}\text{F}$ -FDG PET in the management of cancer. The multidisciplinary panel was initially convened by the American Society of Clinical Oncology with members from the Society of Nuclear Medicine, American College of Radiology, American Cancer Society, Blue Cross and Blue Shield Association (BCBSA), National Coalition of Cancer Survivorship, US Oncology, and American Society for Therapeutic Radiology and Oncology to evaluate the status of the published literature on PET in oncology and to determine whether recommendations on PET could be developed for referring oncology physicians. The SNM subsequently assumed the responsibility for reviewing and evaluating the outcome of the panel's efforts and recommendations. On July 13, 2007, the SNM Board of Directors approved publication of the panel's findings as this special contribution to the *Journal of Nuclear Medicine*.

Most studies that the panel reviewed included PET without CT augmentation. However, the panel realizes PET/CT use is increasingly common and expects PET/CT to further improve the utility of PET.

The use of  $^{18}\text{F}$ -FDG PET in the following types of cancer was assessed: breast, colorectal, esophageal, head and neck, lung, pancreas, and thyroid cancer; lymphoma, melanoma, and sarcoma; and unknown primary tumor. The goal was to provide practitioners with recommendations on the appropriate use of PET in the management of these cancers and to identify gaps in knowledge that may affect future research. Other neoplasms that have been reported and generally recognized as non- $^{18}\text{F}$ -FDG-avid (e.g., renal, prostate, and hepatocellular cancer) were not addressed.

Two principal questions on the appropriateness of  $^{18}\text{F}$ -FDG PET for the management of cancer were addressed: For what cancers should  $^{18}\text{F}$ -FDG PET be used in clinical practice, and under what specific clinical circumstances

should  $^{18}\text{F}$ -FDG PET be used? Recommendations were developed to assist practitioner and patient decisions about health care for specific clinical circumstances (16). It is important to realize, however, that recommendations cannot always account for individual variation among patients. The recommendations are not intended to supplant physician judgment with respect to particular patients or special clinical situations.

## MATERIALS AND METHODS

### Panel Composition

The panel comprised experts in clinical oncology or hematology, radiology or nuclear medicine (specializing in PET), and outcomes or health services researchers with expertise in evidence-based medicine. Both academic and community practitioners were included. A patient representative was also included on the panel.

### Process Overview

In evaluating evidence on the role of PET, the panel was guided by the process established by the GRADE (Grades of Recommendations, Assessment, Development and Evaluation) Working Group (17). This process follows the principle that systematic reviews of the totality of research evidence represent the scientific foundation for development of clinical recommendations (18,19). Therefore, the panel first attempted to identify all systematic reviews on the use of  $^{18}\text{F}$ -FDG PET oncology and used these to assess the quality of primary research evidence (Tables 1 and 2). In doing so, the panel soon clearly saw that the systematic reviews themselves were of varying quality and that a separate assessment of the quality of the systematic reviews was required (Table 3). It also became clear that no systematic review was performed using evidence from randomized controlled trials (RCTs). Because, in general, evidence obtained in RCTs is considered the most reliable (17,20) (Table 4), the panel decided to perform an additional search for randomized evidence and perform its critical appraisal. Therefore, the final recommendations were based on the systematic review of available randomized evidence and an overview (systematic review) of the existing systematic reviews addressing clinical indications of interest (Table 5).

**TABLE 1**  
Definition of Grade of Evidence for Primary Studies and Systematic Reviews

Quality	Definition
High	Further research is unlikely to change confidence in estimate of effect of intervention. No serious limitations were noted.
Moderate	Further research is likely to have important impact on confidence in estimate of effect of intervention and may change estimate. Few serious limitations were noted.
Low	Further research is very likely to have important impact on confidence in estimate of effect of intervention and is likely to change estimate. Typically, more than 2 serious limitations were noted.
Unclear	Any estimate of effect is uncertain. Evidence is either lacking or was not described well enough to allow critical appraisal or make any estimate.

Adopted from recommendations of GRADE Working Group (17).

**TABLE 2**  
GRADE Definitions to Assess Primary Studies

Quality of evidence	Study design	Decrease if . . .	Increase if . . .
High	Randomized trial	Study limitations exist	Association is . . .
Moderate		Serious limitations (−1)	Strong, with no plausible confounders (+1)
Low	Observational study	Very serious limitations (−2)	
Very low	Any other evidence	Important inconsistency is present (−1)	Very strong, with no major threats to validity (+2)
		Directness (generalizability) is uncertain	Evidence exists of a dose–response gradient (+1)
		Some uncertainty (−1)	All plausible confounders would have reduced the effect (+1)
		Major uncertainty (−2)	
		Data are sparse or imprecise (−1)	
		Probability of reporting bias is high (−1)	

### Literature Review and Data Collection

Pertinent systematic reviews and RCTs from the published literature were retrieved and reviewed for the development of these recommendations. Searches of MEDLINE (National Library of Medicine) and other databases (Institute for Clinical Evaluative Sciences, Blue Cross Blue Shield Technology Evaluation Center, and the NHS Health Technology Assessment Program) for pertinent articles were done using strategies developed by Montori et al. (21) and Mijnhout et al. (22). The search was repeated on June 30, 2005, and a final time on March 1, 2006. References from the relevant papers were also searched for further articles of interest. Searches for RCTs were done using the strategy of Haynes and Wilczynski (23). This strategy was combined with the search strategy of Mijnhout et al. (22). In addition, the authors who conducted RCTs were contacted for updated and, in some cases, unpublished information.

### Study Selection

Studies were limited to those in which <sup>18</sup>F-FDG and dedicated PET scanners were used to evaluate any of 11 cancers of interest. The only focus of the report was clinical effectiveness; cost-effectiveness was not considered. Systematic reviews, health technology assessments, and RCTs were eligible.

### Data Extraction

Key evidence from the selected systematic reviews, health technology assessments, and RCTs was extracted using an established hierarchy of diagnostic efficacy (sensitivity, specificity, area under the curve, and summary of receiver operating characteristics) and patient outcomes consequent to performing PET.

### Data Synthesis

For each cancer management decision, the evidence profiles available from each systematic review, health technology assessment, and RCT were generated. A variety of methods was used for the health technology assessments and systematic reviews to summarize evidence of diagnostic efficacy. These ranged from simple summaries of available evidence with no pooling across studies to sophisticated metaanalyses of summary receiver-operating-characteristic curves. The quality of each piece of evidence was assessed. Finally, the authors' conclusions about the sufficiency of all available evidence for each cancer management decision were extracted.

The evidence profiles were distributed to the panel members, who used them during the final panel meeting to make their judgments on the use of PET for each indication.

### Consensus Development Based on Evidence

The entire panel met 3 times. At the first meeting, the panel identified the topics of the recommendations, developed a strategy for completion of the recommendations, and did a preliminary review of the initial literature search. At the second meeting, the panel reviewed the supporting evidence and developed recommendations. Two members of the panel performed the initial grading of the evidence and developed evidence profiles from the eligible studies. These panel members also graded the quality of the systematic reviews and the quality of the primary evidence included in eligible systematic reviews and RCTs and assigned preliminary grades to the overall quality of evidence (Tables 1 and 3). The material was then distributed to the other members of the panel, who had the opportunity to independently verify the content of the evidence profiles and the grades of the quality of evidence. At the third meeting, the panel reviewed the evidence profiles and the quality of evidence. Consensus on the quality of evidence was achieved. Finally, the panel made its recommendations on the use of <sup>18</sup>F-FDG PET for each of these clinical circumstances.

All panel members voted on all aspects of the recommendations (i.e., quality assessment, estimation of benefit or harm, and recom-

**TABLE 3**  
Quality Assessment of Systematic Reviews

Quality	Definition
High	Minimal flaws
Moderate	Minor flaws
Low	Major and extensive flaws
Unclear	Data were described in extremely short form, making it difficult to assess quality of systematic review

Adopted from Oxman's index for assessment of quality of review papers (104).

**TABLE 4**  
Adaptation of GRADE Definitions for Diagnostic Tests

Quality of evidence	Study design	Decrease if . . .	Increase if . . .
High	RCT or prospective cross-sectional (or cohort) studies of patients with diagnostic uncertainty and direct comparison	Study limitations exist* Serious limitations (−1) Very serious limitations (−2) Important inconsistency is present (−1) Directness (generalizability) is uncertain Some uncertainty (−1) Major uncertainty (−2) Data are sparse or imprecise (−1) Probability of reporting bias is high (−1)	Test has low level of false-positive and false-negative results, and good data exist on beneficial consequences of avoiding false-positive and false-negative results (clinical outcomes)
Moderate	Anything else		
Low			
Very low			

\*Such as nonrepresentative population, presence of selection bias, poor choice of gold standard, incomplete performance of tests (verification bias), lack of independent test interpretation (detection bias), and loss to follow-up (attrition bias).

mentations). The coauthors collated all responses, but final agreement on all features of the recommendations was achieved by consensus. The text of the recommendations was repeatedly circulated among all members of the panel until uniform agreement was reached. The only instance of major disagreement occurred when the quality of evidence for the use of PET in head and neck cancer was updated from “poor” to “moderate” after a member of the panel challenged the initial appraisal of primary research evidence. After the external review, the quality of evidence on the role of PET in the setting of solitary pulmonary nodule (SPN) was downgraded to “moderate” from “high.” There were no instances of disagreement in the actual recommendations.

These recommendations were circulated in a draft form, and all members of the panel had a further opportunity to comment on the strength of the recommendations, the quality of the evidence, and the systematic grading of the data supporting each recommendation. Final text editing was performed by 2 members of the panel.

The panel did not consider the cost-effectiveness of <sup>18</sup>F-FDG PET.

### Accuracy Versus Impact on Decision Making

In evaluating the usefulness of PET, the panel’s original intent was to incorporate the consequences of a decision of ordering PET and the effect of the subsequent PET results on patient outcomes. Unfortunately, most of the literature has focused on evaluating the accuracy of PET (as expressed by calculation of sensitivity, specificity, and area under the curve) instead of on evaluating the clinical value of PET information on decision making. This situation is not unique to PET but plagues the entire state of the evaluative science of diagnostic testing. Hilden, for example, wrote about the schism between the “ROCgraphers,” who are interested solely in test accuracy (receiver operating characteristic curve [ROC]), and the “VOIgraphers,” who are interested in the clinical value of information (VOI) associated with the consequences of decision making (24). Consequently, the desired level of details on the clinical consequences of ordering PET is extremely poor. In addition, the positive predictive values (PPVs) and negative predictive values (NPVs) of PET, as indeed those of any other diagnostic test, are a function not only of test accuracy (as expressed in terms of test

sensitivity, specificity, or likelihood ratio) but also of the prevalence (prior probability) of a condition in the population of patients on whom PET is performed (25–27). Because precise estimates of the prevalence of cancers were difficult to deduce from the studies the panel evaluated, caution is urged in translating predictive values from the numbers quoted in these recommendations to actual practice. Data on prevalence to allow calculation of PPVs or NPVs were typically not reported. The same applied to the issue of harm potentially associated with performing PET. Often, it was not clear whether this lack of reporting on potential harm was because no harm was associated with PET or because data on harm were not collected. Similarly, data were not reported in a standardized fashion (28); for example, although some authors reported sensitivity or specificity, others reported likelihood ratios, areas under the curve, or diagnostic odds ratios, and a few reported change-in-management outcomes. Finally, because of the enormity of the task, the panel found it impossible to link its recommendations on PET to specific treatment recommendations to provide precise quantitative thresholds above which particular clinical conditions of interest (e.g.,

**TABLE 5**  
Overall Grade of Evidence: Quality of Primary Research plus Quality of Systematic Reviews

Grade	Definition
High-quality evidence	High-quality systematic reviews with high-quality evidence and consistent findings
Moderate-quality evidence	High-quality or moderate-quality systematic reviews with moderate-quality evidence and consistent findings
Low-quality evidence	High-, moderate-, low-, or unclear-quality systematic reviews with low-quality or unclear evidence and consistent or inconsistent findings
	Moderate-quality systematic reviews with moderate-quality evidence and inconsistent findings

assessment of the stage of disease) could be considered ruled in or ruled out. For this reason, all judgments made by the panel were qualitative.

### Why PET Rather Than PET/CT?

The current recommendations concern only dedicated PET systems in the management of patients with cancer and do not consider the role of dual-modality PET/CT systems. The primary reason for the exclusion of PET/CT relates to the paucity of published data on PET/CT and the decision of the GRADE Working Group to specifically exclude commentary on dual-modality systems at the time the group developed its process. It is clear from published individual articles that the combined, concurrent use of PET and CT in dual-modality systems provides a diagnostic advantage over either PET alone or CT alone (29–33). This advantage is not surprising, because the near-simultaneous acquisition of superb CT anatomic information with PET biochemical characterization allows for precise anatomic localization of  $^{18}\text{F}$ -FDG PET abnormalities and accurate characterization of the metabolic activity of indeterminate CT masses. Commercially available systems can provide a high-quality diagnostic-level contrast-enhanced CT study along with the  $^{18}\text{F}$ -FDG PET study and whole-body torso image (base of brain to mid thigh) in 30 min or less. Clinical sites offering PET/CT, and clinicians selecting its use for their patients with cancer, can expect enhanced diagnostic performance over dedicated PET or CT. However, only a few well-done studies addressing the issue of dual-modality PET/CT systems exist in the literature, and no systematic review has been published on the topic. The issue of PET/CT will be addressed when more mature data become available (likely within 2–3 y).

### Summary of Outcomes Assessed

In this article, the panel presents recommendations on using  $^{18}\text{F}$ -FDG PET for different types of cancer: breast, colorectal, esophageal, head and neck, lung, thyroid, and pancreatic cancer; melanoma, sarcoma, and lymphoma; and unknown primary tumor. The use of  $^{18}\text{F}$ -FDG PET is presented in the context of diagnosis, staging, and detection of the recurrence or progression of a given malignancy (Table 6). The panel decided not to evaluate the use of  $^{18}\text{F}$ -FDG PET in assessing response to treatment. That topic will be addressed in the future.

## RESULTS

Supplemental tables containing complete evidence-summary data are available online at <http://jnm.snmjournals.org>.

Ideally, the optimal way to evaluate the accuracy and clinical impact of a diagnostic test is to conduct an RCT (20). This undertaking is, however, logistically and often ethically difficult. The second best way is through consecutive, prospective enrollment of adequate numbers of patients, all of whom have undergone both index and gold standard testing with masked interpretation of the index test. The panel's systematic review of the existing evidence proved that few high-quality studies evaluating the role of  $^{18}\text{F}$ -FDG PET in oncology have been performed. Fifty-two systematic reviews were initially identified from the liter-

ature survey; 36 systematic reviews were eligible for this overview. In addition, the panel identified 3 RCTs (Fig. 1).

The panel found no systematic review or RCT using the newer technology (i.e., integrated PET/CT scanners). The quality of the systematic reviews varied considerably. Existing systematic reviews often combined retrospective and prospective data without sensitivity analysis according to the quality. Often, it was not clear whether the evidence on a given topic was missing or was not reported. Systematic reviews frequently did not provide basic information such as the number of prospective or retrospective studies, the number of patients, or the type of PET scanner used. Most studies suffered from spectrum bias (i.e., they included a population consisting of the entire clinical spectrum with early- and late-stage cancer), selection bias (i.e., they ordered tests on the basis of certain patient characteristics), verification bias (i.e., they administered the reference standard test only to patients with a positive PET result), and detection bias (i.e., they failed to perform masked interpretation of PET images). However, some members of the panel believed that although it is important to address the problem of detection bias in research studies, in clinical practice PET studies are always interpreted in conjunction with other available imaging studies. Future research should define the extent to which masked reading of PET studies with or without other imaging studies affects the accuracy of PET interpretation.

Furthermore, the PET procedure and eligible patient populations were often poorly described (e.g., details on the number of patients with diabetes, on glucose level, on the use of contrast material, on attenuation correction, and on reproducibility were typically not reported). Investigators generally regarded concordant findings between PET and CT or PET and MRI as either true-positive or true-negative, failing to appreciate that both PET and CT or PET and MRI can concordantly give false results. Therefore, it is fair to say that, in general, evidence on the use of PET in oncology is far from perfect.

Nevertheless, some patterns have emerged from the panel's systematic assessment of evidence. With a few exceptions, conventional imaging tests (e.g., ultrasonography, CT, and bone scanning) have rarely been superior to PET in any clinical indication. In general, NPV is greater than PPV, although NPV and PPV depend critically on the prevalence of disease in the population or the pretest probability in a given patient. Any positive finding should be confirmed by subsequent histopathologic examination because of the considerable risk of false-positive results. In general, the panel concluded that  $^{18}\text{F}$ -FDG PET should be used as an imaging tool additional to conventional radiologic methods such as CT or MRI.

## BREAST CANCER

### Is $^{18}\text{F}$ -FDG PET Useful for Differentiating Cancer from Benign Mammographic Lesions?

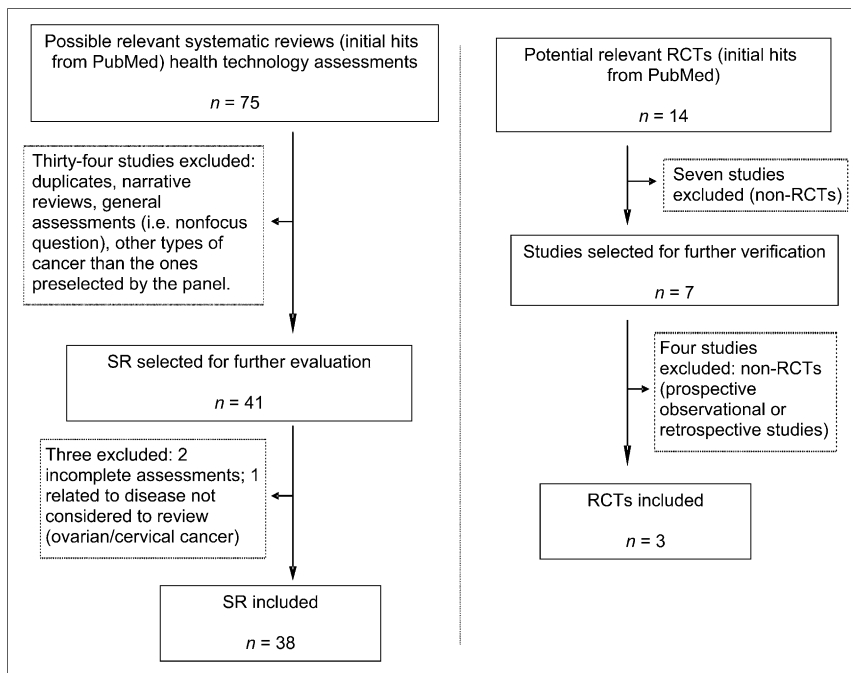
*Background and Rationale.* Screening for breast cancer and detection of early stages of tumors are believed to have

**TABLE 6**  
Summary of Recommendations

Disease	Objective	Recommended?	Net benefits?	Overall quality
<b>Breast cancer</b>				
Diagnosis		No	Uncertain	Moderate
Staging	Local (axillary lymph node) staging	No	No	Moderate
Recurrence	Detection of metastases	Yes, for patients clinically suspected of metastasis/recurrence	Yes	Moderate
<b>Colorectal cancer</b>				
Diagnosis		No	No	Low
Staging	Liver metastasis staging	Yes	Yes	Moderate
Recurrence	Detection of metastases or local recurrence	Yes	Yes	Moderate
<b>Esophageal cancer</b>				
Diagnosis		No*		
Staging	Pretreatment staging	Yes, especially to detect distant metastases	Yes	Moderate
Recurrence		No*		
<b>Head and neck cancer</b>				
Diagnosis	Diagnosis of unknown primary tumor	Yes	Yes	Moderate
	Diagnosis of head and neck tumors	No	Uncertain	Low
Staging		Yes	Yes	Moderate
Recurrence		Yes	Yes	Moderate
<b>Lymphoma</b>				
Diagnosis		No*		
Staging	General staging Bone marrow staging	Suggested Yes	Yes Yes	Low Moderate
Recurrence <sup>†</sup>		Yes, for HD and NHL after completion of initial treatment	Yes	Moderate
		No, for general follow-up of asymptomatic HD or NHL	No	Low
<b>Lung cancer</b>				
Diagnosis		SNP: yes	Yes	Moderate
Staging		NCLC: yes SCLC: no recommendation is possible	Yes Uncertain	High Low
Recurrence		No*		
<b>Melanoma</b>				
Diagnosis		No*		
Staging		Yes, for distant metastases	Yes	Moderate
Recurrence		No*		
<b>Pancreatic cancer</b>				
Diagnosis		Yes, for cases in which CT is inconclusive	Yes	Moderate
Staging		No*		
Recurrence		No*		
<b>Sarcoma</b>				
Diagnosis		No recommendation is possible	Uncertain	Low
Staging		No recommendation is possible	Uncertain	Low
Recurrence		No*		
<b>Thyroid cancer</b>				
Diagnosis		No*		
Staging		No*		
Recurrence		Yes, if <sup>131</sup> I whole-body findings are negative and thyroglobulin is elevated	Yes	Low
		No, if <sup>131</sup> I whole-body findings are negative and thyroglobulin is normal (surveillance)	No	Low
<b>Unknown primary tumor</b>				
Diagnosis		Yes	Yes	Low
Staging		No*		
Recurrence		No*		

\*Synthesis of research data has not been performed.

<sup>†</sup>Restaging/detection of relapse; assessment of residual mass or progression after completion of initial treatment.



**FIGURE 1.** PET in oncology: flowchart for selection of systematic reviews and RCTs. Last search update was conducted Nov. 1, 2006, on PubMed using strategy of Montori et al. (27).

led to a decrease in breast cancer mortality (34). Mammography is the principal imaging tool to screen for breast cancer. However, most abnormalities detected by screening mammography are benign on biopsy (35). Consequently, patients may experience unnecessary harm. The use of  $^{18}\text{F}$ -FDG PET in screening for and diagnosing primary breast cancer could help decrease unnecessary biopsies.

**Evidence Summary.** The panel concluded against the routine use of  $^{18}\text{F}$ -FDG PET in diagnosing breast cancer. The panel found moderate evidence against routine use and concluded that the possibility of missing early-stage lesions and the high risk of false-negative results may be detrimental. However, the panel also indicated that in specific clinical circumstances and selective cases (e.g., high-risk patients with masses > 2 cm or aggressive malignancy and serum tumor marker elevation), physicians might decide to modify this recommendation (36).

The panel also concluded that imaging with  $^{18}\text{F}$ -FDG PET is not useful for screening purposes.

**Status of the Evidence.** The panel identified 3 systematic reviews (37–39) that addressed this issue. Facey et al. (38) addressed the use of  $^{18}\text{F}$ -FDG PET in the differential diagnosis of benign from malignant lesions. Facey et al. found 5 studies (14–144 patients per study) that were conducted before 2004. The overall quality of evidence in this systematic review was low, and the quality of the single systematic review itself (38) was unclear because it was presented as a short report. The BCBSA systematic review (37) used a quality index score, partially validated, to assess the quality of primary research studies. The quality of primary research evidence was moderate. No randomized studies were performed on the topic, and the reviews did not clearly state how many studies were retrospective and how many were prospective and enrolling

consecutive patients. None of the studies met all quality criteria. The major deficiencies were a failure to adequately describe the PET procedure (including a reference to glucose level and the method of interpretation), the work-up for resection, the inclusion and exclusion criteria for patients, and the impact on patient management. The authors found that, in all 5 studies,  $^{18}\text{F}$ -FDG PET had sensitivity and specificity greater than 80% and 76%, respectively.  $^{18}\text{F}$ -FDG PET results were comparable to mammography results in 2 studies.

The reviews of the BCBSA (37) and of Facey et al. (38) also addressed questions about patients with breast masses or abnormal mammography findings and negative  $^{18}\text{F}$ -FDG PET findings who underwent biopsy. Thirteen studies were found in the BCBSA review (37), with 16–144 patients per study. The pooled metaanalysis of 10 studies showed that PET had a sensitivity of 89% (95% confidence interval [CI], 84%–93%) and a specificity of 80% (95% CI, 70%–87%). NPV was 88% if the projected prevalence of breast cancer was 50%. The calculated positive likelihood ratio (LR+) and negative likelihood ratio (LR–) were 4.45 and 0.14.

The Agency for Healthcare Quality and Research (39) published a comparative effectiveness review and evaluated different imaging tests for the diagnosis of breast abnormalities, including 9 studies on the use of PET. This report found that in suggestive lesions in general, the sensitivity of PET was 82%, compared with 93% for MRI and 86% for ultrasonography. The specificity for PET was 78%, versus 72% for MRI and 66% for ultrasonography. The NPV at a 20% prevalence of breast cancer was 92% for PET, compared with 96% for MRI and 95% for ultrasonography. Similar estimates were not calculated for mammography. The calculated LR+ and reported LR– for PET were 3.78 and 0.33, respectively.



The overall quality of evidence was moderate, the quality of the systematic review itself was high, and the quality of primary research evidence was moderate. Once again, no randomized studies were performed. Many studies suffered from verification bias and spectrum bias (i.e., enrolled patients having different stages of disease), and a few were retrospective. It was also not clear how many overlapping studies were included in these 2 different reviews. However, the findings across the reviews and studies themselves were consistent.

The panel concluded that the use of  $^{18}\text{F}$ -FDG PET for the diagnosis of primary breast cancer may not be beneficial because of the possibility of missing early-stage lesions and the high risk of false-negative results. The panel also determined that MRI and screening mammography appear to be superior to PET.

*Areas of Uncertainty and Implications for Future Research.* Although other competing imaging modalities appear superior to PET, the quality of evidence for using PET in this setting is moderate, and further research may therefore alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use. The panel believes that future research should focus on the role of the dedicated breast PET scanners that likely will be needed to provide a sufficiently competitive spatial resolution for differential diagnosis of small (clinically relevant) breast lesions.

#### **Is $^{18}\text{F}$ -FDG PET Useful for Assessing Axillary Involvement in Breast Cancer Patients?**

*Background and Rationale.* The management and prognosis of breast cancer depends on a variety of factors, including the complete staging of the tumor, locally and distantly. The detection of axillary lymph nodes and distant metastases is extremely important in deciding how to treat breast cancer patients.

*Evidence Summary.* The panel concluded against routine use of  $^{18}\text{F}$ -FDG PET for axillary staging of breast cancer. The panel found moderate evidence that the use of PET will likely misclassify the extent of breast cancer and concluded that  $^{18}\text{F}$ -FDG PET may not be beneficial, mostly by failure to lead to appropriate treatments.

*Status of the Evidence.* The panel identified 2 systematic reviews (37,38), which analyzed 10 studies enrolling 18–167 patients per study. Whether there was an overlap in studies between the reviews was difficult to verify. The overall quality of evidence was moderate, the quality of the systematic reviews themselves was high, and the quality of primary research evidence was moderate. The primary research evidence might have suffered from many biases, including spectrum and verification biases. There was also a mix of data, including patients and lesions. Of the 10 studies, only 1 was classified as of high quality, but most studies were consistent in their findings. Nine studies had available data on sensitivity and specificity; in 7 studies PET had a sensitivity of at least 85%, and in 6 studies PET had a specificity of at least 90%. Only 4 studies analyzed

patients with no palpable axillary lymph nodes, with a total of 203 patients, and the pooled metaanalyses showed a sensitivity of 80% (95% CI, 46%–95%) and a specificity of 89% (95% CI, 83%–94%). The calculated LR+ and LR– of PET were 7.27 and 0.22, respectively.

$^{18}\text{F}$ -FDG PET was also compared with axillary lymph node dissection (ALND) or ALND plus sentinel node biopsy (SNB). The overall quality of evidence was moderate, the quality of the systematic review itself was high, and the quality of primary research evidence was moderate. The evidence reviews suffered from many biases, including possible verification, spectrum, and detection biases. Eight studies were identified; most were prospective, and they included 15–129 patients per study. Only 1 high-quality study was found, and overall, the risk of false-negative results was high. If ALND was used as the reference test, PET showed a sensitivity of 40%–93% and a specificity of 87%–100%. If ALND + SNB was used as the reference test, PET showed a sensitivity of 68%–96% and a specificity of 57%–80%. The prevalence of node-positive disease was 33%–64%. PET accuracy was lower when evaluated against ALND + SNB. Assuming that PET findings were negatives, the NPV of PET was 92.1%, given a prevalence for node-positive disease of 30%.

The panel concluded that the use of PET in staging axillary lymph nodes for breast cancer is not beneficial. There is indirect evidence that PET might falsely understage disease in patients with earlier stages of breast cancer, resulting in failure to lead to appropriate treatment.

The panel also determined that PET is inferior to axillary node dissection and sentinel node biopsy in this setting.

*Areas of Uncertainty and Implications for Future Research.* The amount of evidence is sufficient to conclude against the use of  $^{18}\text{F}$ -FDG PET for assessing axillary involvement with breast cancer. SNB is well established as a diagnostic standard. It is highly unlikely that future PET studies, even if they are methodologically better than the available ones, will change this conclusion.

#### **Is $^{18}\text{F}$ -FDG PET Useful for Detecting Metastatic or Recurrent Breast Cancer?**

*Background and Rationale.* The background and rationale are the same as those described for the use of  $^{18}\text{F}$ -FDG PET for assessing axillary involvement in breast cancer.

*Evidence Summary.* The panel concluded that  $^{18}\text{F}$ -FDG PET should routinely be added to the conventional work-up in detecting metastatic or recurrent breast cancer in those patients clinically suspected of metastasis or recurrence. The panel found moderate evidence that  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries. The panel did not find PET useful for surveillance of patients who are asymptomatic.

*Status of the Evidence.* Isasi et al. (40) performed a systematic review and metaanalysis to address the impact of  $^{18}\text{F}$ -FDG PET in detecting metastatic disease and recur-

rence. The authors included articles published up to June 2004. The overall quality of evidence was moderate, and the quality of the systematic review itself was high. The systematic review used a quality index score to assess the primary research studies. The quality of primary research evidence was moderate. The review identified 18 studies, with 18–75 patients per study. <sup>18</sup>F-FDG PET was performed to evaluate clinically suspected recurrences or metastases of breast cancer in 14 studies. The studies used pathology or clinical follow-up as the reference standard. Seven studies were retrospective, 6 were prospective, and 5 had an unclear design. Eight studies reported that the interpreters of PET images were masked to the reference test, 2 reported that interpretation was not masked, and 6 did not mention whether those who interpreted the PET images were aware of the results of the reference test. The major deficiencies were spectrum and verification biases. The studies using patient-based data resulted in a pooled sensitivity of 90%, and the specificity was 87%. The calculated LR+ and LR– of PET were 6.92 and 0.11, respectively. The maximum joint sensitivity and specificity was 86%. The studies using lesion-based data had a median sensitivity of 92% (range, 57%–97%) and a median specificity of 89% (range, 79%–96%). The pooled true-positive rate was 85% and false-positive rate was 7%. The maximum joint sensitivity and specificity was 89%.

The panel concluded that, overall, using <sup>18</sup>F-FDG PET to restage or evaluate for recurrence of breast cancer in those patients clinically suspected of metastasis or recurrence is beneficial. However, <sup>18</sup>F-FDG PET should not be used to replace CT but should complement the current approach to the diagnostic work-up.

*Areas of Uncertainty and Implications for Future Research.* The evidence is sufficient to support the use of <sup>18</sup>F-FDG PET for restaging or recurrence. Although the quality of evidence is moderate, future research might better clarify the role of PET, particularly in the setting of surveillance of asymptomatic patients.

## COLORECTAL CARCINOMA

### Is <sup>18</sup>F-FDG PET Useful for Diagnosing Colorectal Carcinoma?

*Background and Rationale.* Colorectal cancer is the second most common cause of cancer death in the United States (41). Any method that can help detect colorectal cancer early in its course will likely be useful. Whether <sup>18</sup>F-FDG PET has a potential utility in early detection of colorectal cancer is not sufficiently known.

*Evidence Summary.* The panel recommended against the routine use of <sup>18</sup>F-FDG PET for detecting primary colorectal carcinoma. The panel found little evidence to support the use of <sup>18</sup>F-FDG PET for this indication.

*Status of the Evidence.* The panel found 1 systematic review (38) that evaluated the detection of malignant primary tumor and was conducted before 2004. This review was briefly reported and evaluated only 2 studies enrolling 16 and 24 patients. The overall quality of the evidence was

low, the quality of the systematic review itself was unclear, and the quality of primary research evidence was low. <sup>18</sup>F-FDG PET sensitivity was 85% in both studies, but specificity was reported only in 1 study and was 67%.

The panel concluded that, overall, using <sup>18</sup>F-FDG PET in the primary diagnosis of colorectal cancer is not beneficial.

*Areas of Uncertainty and Implications for Future Research.* The existing evidence is extremely limited and too low to justify any recommendation. High-quality studies are needed in this area before <sup>18</sup>F-FDG PET can be considered for the diagnosis of primary colorectal tumor.

### Is <sup>18</sup>F-FDG PET Useful for Managing Colorectal Liver Metastasis?

*Background and Rationale.* Liver metastasis is the main cause of death in patients with colorectal cancer. However, for selected patients in whom recurrent disease is confined to the liver, surgical resection of the metastases may be curative, with a 5-y survival of greater than 30% (42). Conventional imaging with CT often fails to identify preoperatively those patients whose metastases can successfully be resected: About 15%–25% of cases are deemed unresectable at the time of surgery, and cancer recurs within 3 y in 60% of patients whose disease was deemed to be resectable. Therefore, having better imaging techniques to improve staging and to avoid futile surgery is clearly desirable.

*Evidence Summary.* The panel concluded that <sup>18</sup>F-FDG PET should be used routinely in addition to conventional imaging in the preoperative diagnostic work-up of patients with potentially resectable hepatic metastases from colorectal cancer. The panel found moderate evidence that the use of PET will likely improve important health-care outcomes and concluded that PET is beneficial, mostly by avoiding futile surgeries.

*Status of the Evidence.* The panel identified 6 systematic reviews on this topic (38,43–47), which analyzed 8–61 studies enrolling up to 145 patients per study. The largest systematic review (45) included 1,058 patients who were studied by PET. An overlap of studies among the systematic reviews was possible. One systematic review included studies up to 1999, 2 systematic reviews included studies up to 2000, 1 included studies up to December 2003, and 2 included studies up to 2004. The quality of primary research evidence was moderate to unclear. Although the quality of the systematic reviews themselves was unclear to high, the overall quality of the evidence was moderate. No randomized studies were performed, and the reviews did not make clear how many studies were retrospective and how many were prospective and enrolling consecutive patients. The major deficiencies were possible selection, verification, and spectrum biases. The sensitivity and specificity of PET for hepatic lesions were higher than 85% in most studies. The 2 most recent systematic reviews (45,46) reported sensitivities per patient of 95% (95% CI, 93%–96%) and had a higher sensitivity than did CT and MRI. Wiering et al. (46) identified 32 studies: PET had a sensitivity for hepatic lesions of 88%

(95% CI, 88%–98%) and a specificity of 96% (95% CI, 70%–100%), and CT had a sensitivity of 83% (95% CI, 64%–89%) and a specificity of 84% (95% CI, 68%–97%). The calculated LR+ and LR– of PET were 22 and 0.12, respectively. In this review, 6 studies were classified as having highest quality scores. For these studies, the authors reported that PET had a sensitivity for hepatic lesions of 80% and a specificity of 92% and CT had a sensitivity of 86% and a specificity of 88%. In this subgroup, the calculated LR+ and LR– of PET were 10 and 0.22, respectively. In these 6 studies, the mean change-in-management rate was 25% (range, 20%–32%) (46).

Bipat et al. (45) demonstrated that when sensitivity was calculated on a per-lesion basis, PET sensitivity (76%) was comparable with helical CT sensitivity (64%) and with 1.0-T and 1.5-T MRI sensitivity (66% and 64%, respectively); nonhelical CT had a lower sensitivity (52%).

The BCBSA review (43) addressed the impact of a change in management brought about through PET findings. They found that 11 studies (680 patients) reported on the proportion of patients for whom PET affected management decisions (not exclusively assessing hepatic metastases only). The range of change in management was 7%–68% (average, 20%). PET was influential in ruling out (unnecessary) surgery in 12% and influenced initiating surgery in 8%. The decision to avoid (unnecessary) surgery in 60% of patients was affected by PET results (43).

The panel concluded that, overall, using <sup>18</sup>F-FDG PET is beneficial in addition to CT in the preoperative diagnostic work-up of patients with potentially resectable hepatic metastases of colorectal cancer or in evaluations for recurrence. The greatest benefit was attributed to avoidable futile surgeries and help with determining appropriate treatment.

*Areas of Uncertainty and Implications for Future Research.* Despite the moderate to unclear quality of the available evidence, the consistency in the findings indicated that PET has clinical utility in the detection of hepatic metastases. Nevertheless, definitive high-quality studies could be useful—particularly contrast-enhanced PET/CT studies, which the panel believes to be the most promising imaging modality in this setting.

### Is <sup>18</sup>F-FDG PET Useful for Detecting Extrahepatic Recurrence or Local Relapse?

*Background and Rationale.* Extrahepatic metastatic disease is considered incurable. An accurate imaging modality to assess whether colorectal cancer has spread beyond the liver may help to avoid unnecessary surgeries. In addition, many patients treated with colorectal surgery commonly present with lesions in the colon, which could represent a local recurrence or postoperative scarring. Differentiating between these 2 conditions would obviously be useful to these patients.

*Evidence Summary.* The panel concluded that <sup>18</sup>F-FDG PET should routinely be obtained after the conventional work-up, especially if carcinoembryonic antigen levels are

increased and the results of the conventional work-up are negative. PET can also be used to differentiate between local relapse and postsurgical scars, but no evidence is available to help to define the timing and sequence of PET in relationship to other imaging techniques. The panel found moderate evidence that the use of PET will likely improve important health-care outcomes and concluded that the use of <sup>18</sup>F-FDG PET is beneficial, mostly by avoiding futile surgeries. Depending on clinical circumstances, physicians may decide to modify this recommendation.

*Status of the Evidence.* The panel identified 3 systematic reviews that studied the use of PET for detecting extrahepatic lesions (38,46,47). These systematic reviews analyzed 5 and 32 studies conducted up to 1999 and 2004, respectively. The overall quality of evidence was low to moderate, the quality of the systematic reviews themselves was rated from unclear to high, and the quality of primary research evidence was moderate. One systematic review (47) used a quality index score to assess the primary research studies. Only 3 of the 11 studies met more than 75% of the criteria stipulated by the reviewers. No randomized studies were performed, and the review did not make clear how many studies were retrospective and how many were prospective and enrolling consecutive patients. For extrahepatic lesions, the whole-body sensitivity and specificity of PET were 97% and 76%, respectively. The calculated LR+ and LR– of PET were 4.04 and 0.04, respectively. The effect of PET on management (a change in management when the PET findings were ultimately correct), based on data pooled from 7 studies, was 29% (range, 20%–44%). For example, some patients were able to avoid unnecessary surgery as a result of the PET findings (47).

The most recently published systematic review (46) reported that for extrahepatic lesions, the sensitivity and specificity of <sup>18</sup>F-FDG PET were 92% and 95%, respectively. The calculated LR+ and LR– of PET were 18.40 and 0.08, respectively. The sensitivity and specificity of CT for extrahepatic lesions were 61% and 91%, respectively. The results were consistent across all primary studies. When data were pooled from the 6 studies that had the highest quality scores, the sensitivity and specificity of PET for extrahepatic lesions were 91% and 98%, respectively. For CT, sensitivity and specificity were 55% and 96%, respectively. PET resulted in a change in clinical management 32% (range, 20%–58%) of the time in 13 of 17 studies with quality scores above the mean. In the 6 studies with the highest quality scores, the mean change in management was 25% (range, 20%–32%) (46).

In addition, the panel found a review that assessed the role of PET in differentiating local recurrence from postoperative scarring (43). Six studies were included, enrolling 198 patients. The overall quality of evidence was moderate, the quality of the systematic review itself was high, and the quality of primary research evidence was moderate. Only 1 study clearly stated that histopathology was used as the reference test, and whether the interpretation had been

masked was not clear for all studies. The overall sensitivity was 96%, and the specificity was 98%. The calculated LR+ and LR- of PET were 48 and 0.04, respectively. In these studies, the pooled prevalence of malignancy was 69%, and the NPV was 92%. When the prevalence of malignancy was only 5%, the NPV was 99.8%.

The panel concluded that, overall, the use of PET in addition to CT is beneficial for evaluating recurrence of colorectal carcinoma if CT is inconclusive, if carcinoembryonic antigen levels are increased, or if local relapse is clinically suspected. Most of the benefit is attributable to detection of extrahepatic metastases, which generally preclude liver resection. This detection, in turn, should help to avoid futile surgeries and to determine appropriate treatment.

*Areas of Uncertainty and Implications for Future Research.* Despite the variation in the quality of available evidence, the reported findings were in general consistent, indicating that PET has clinical utility for the staging or restaging of extrahepatic metastases or locoregional recurrence. Nevertheless, better-designed studies will be useful for definitive assessment of the role of <sup>18</sup>F-FDG PET in detecting recurrence of colorectal cancer.

## ESOPHAGEAL CANCER

### Is <sup>18</sup>F-FDG PET Useful for Staging Esophageal Cancer?

*Background and Rationale.* For esophageal cancer, surgery offers the best chance of cure. The poor long-term survival of patients who have a complete tumor resection seems to be related to failure to detect distant metastases at the time of surgery (48). The survival rates of noncurative surgery are similar to those of nonsurgical therapy using combined chemoradiation (49). Surgery in patients with advanced disease can be avoided if accurate preoperative staging information is available. Today's stage-adjusted treatment of advanced esophageal cancers requires a meticulous diagnostic work-up including the use of standard staging tools (endoscopy, endoscopic ultrasonography, or CT) (50). However, recent studies have reported efficacy for PET suggesting that it is more accurate than conventional imaging modalities in this setting (51).

*Evidence Summary.* The panel concluded that PET should routinely be used as an additional tool for staging esophageal cancer. The panel found moderate evidence that the use of <sup>18</sup>F-FDG PET will likely improve important health-care outcomes and concluded that the use of <sup>18</sup>F-FDG PET is beneficial, mostly by avoiding futile surgeries.

*Status of the Evidence.* The panel identified 2 systematic reviews (38,52) that addressed the use of <sup>18</sup>F-FDG PET for staging esophageal cancer. The overall quality of evidence was moderate, the quality of the systematic reviews themselves ranged from low to high, and the quality of primary research evidence was moderate. A well-done systematic review of the use of <sup>18</sup>F-FDG PET in preoperative staging of esophageal cancer evaluated 12 studies up to June 2003 with

a total 490 patients (52). The primary evidence suffered from spectrum bias (4 studies) and verification bias (3 studies). Ninety-two percent of studies did not describe whether the PET interpretation was masked. For detection of local nodal metastases, PET had a sensitivity and specificity of 51% (95% CI, 34%–69%) and 84% (95% CI, 76%–91%), respectively. The calculated LR+ and LR- of PET were 3.19 and 0.58, respectively. The mean prevalence was 55%, with a PPV of 60% and an NPV of 46%. The available data suggest that in the detection of locoregional disease, PET appears to be inferior to endoscopic ultrasonography. For the detection of distant metastases, PET had a sensitivity and specificity of 67% (95% CI, 58%–76%) and 97% (95% CI, 90%–100%), respectively. The calculated LR+ and LR- of PET were 22.3 and 0.34, respectively. The mean prevalence was 36%, with a PPV of 92% and an NPV of 83%.

The panel judged that, overall, the use of <sup>18</sup>F-FDG PET is beneficial for staging esophageal cancer. <sup>18</sup>F-FDG PET can be particularly useful as an additional tool for the detection of distant metastases; however, its accuracy for the detection of local nodal metastases is still modest. On balance, evidence supports the use of <sup>18</sup>F-FDG PET in preoperative staging of esophageal cancer.

*Areas of Uncertainty and Implications for Future Research.* The evidence is sufficient to support the use of <sup>18</sup>F-FDG PET for detecting distant metastases. Detection of local nodal metastases was less accurate. However, because of clinical heterogeneity across studies that evaluated PET for detecting local node disease, future research may change this finding.

## HEAD AND NECK CANCER

### Is <sup>18</sup>F-FDG PET Useful for Detecting Clinically Suspected Unknown Head and Neck Primary Tumors?

*Background and Rationale.* The incidence of unknown primary tumors in the head and neck region is 3%–7% of all head and neck cancers (53). The traditional evaluation consists of a careful ambulatory examination, including fiberoptic laryngoscopy or nasopharyngoscopy, detailed imaging (typically CT or MRI), and panendoscopy, with directed biopsies of at-risk sites and tonsillectomy. Recently, <sup>18</sup>F-FDG PET has been demonstrated to be a useful diagnostic imaging study in this situation (54).

*Evidence Summary.* The panel concluded that PET should be added to the imaging tests routinely used to identify unknown primary head and neck tumors. The panel found moderate-quality evidence that the addition of <sup>18</sup>F-FDG PET will likely improve important health-care outcomes and concluded that <sup>18</sup>F-FDG PET is beneficial in this context. However, regardless of whether the initial PET findings are negative or positive, biopsy should be performed. PET would not be considered superfluous because when its findings are negative it should be followed by multiple masked biopsies whereas when its findings are positive it will direct biopsy toward a PET-positive lesion.

*Status of the Evidence.* The panel identified 4 systematic reviews (38,55–57), which analyzed 7 or 8 studies each. The latest systematic review searched studies up to February 2003 (57). The panel could not exclude the possibility of overlap between studies in the systematic reviews. The overall quality of the evidence was moderate, the quality of the systematic reviews varied from low to high, and in all systematic reviews the quality of some primary research studies was low because of problems with verification, detection, or spectrum bias. However, the quality of many other individual research studies, particularly those recently published, was high. In these studies, all patients (with rare exceptions in individual studies) had biopsy verification of disease and often multiple biopsy sampling procedures to exclude other sites in the head and neck. The patients were consistent in the papers published, with all patients having undergone standard clinical staging evaluations including direct panendoscopy either before or after PET.

PET was able to identify the unknown primary tumor in more than 20% of patients when clinical findings were negative. Nieder et al. (56) reported an overall sensitivity of 67%, specificity of 82%, PPV of 56%, and NPV of 86%. The calculated LR+ and LR– were 3.72 and 0.40, respectively.

The panel concluded that the use of PET is beneficial if conventional imaging findings are negative in identifying unknown primary tumors of the head and neck. However, if the PET findings are negative, further effort should be made to identify the primary tumor because of the chance of false-negative results. If the PET findings are positive, confirmatory biopsy is necessary because of the risk of false-positive results. The value of performing panendoscopic evaluation before or after PET and the extent of the evaluation are issues that may need to be decided on an individual basis because 10% of the patients in some reports (58–61) had primary tumors in other areas (e.g., lung or esophagus) that would obviate a further work-up for head and neck cancer.

*Areas of Uncertainty and Implications for Future Research.* Despite the consistency in available evidence showing a role for <sup>18</sup>F-FDG PET as an adjuvant tool to detect head and neck tumors presenting as an unknown primary, PET is still not sufficiently accurate to replace panendoscopy. Further research using the improved technique of PET/CT may alter confidence in the effect of <sup>18</sup>F-FDG PET for this use and may change the estimate. High-quality studies in this clinical setting are needed.

### **Is <sup>18</sup>F-FDG PET Useful for Diagnosing Head and Neck Tumors?**

*Background and Rationale.* Cancer of the head and neck affects 30,000 Americans annually. Recurrent and systemic diseases usually are unresponsive to therapy. Early diagnosis and accurate staging are essential for treatment and for improving the prognosis (62). The standard imaging tests for the diagnosis of head and neck tumors, CT and MRI, are far from perfect. Better imaging methods will be useful.

*Evidence Summary.* The panel concluded against routine use of PET in addition to CT or MRI in the diagnostic work-up of primary-tumor head and neck malignancies. The quality of evidence is insufficient to allow a confident judgment on whether PET can determine the anatomic extent of primary head and neck malignancies at the level of certainty required for surgical resection.

*Status of the Evidence.* The panel identified 1 systematic review (57), which analyzed 4 studies that compared PET with CT or MRI to characterize squamous cell carcinoma. The overall quality of the evidence was low, the quality of the systematic review itself was low, and the quality of primary research evidence was unclear. No randomized studies were performed, and the sample sizes of the studies were not described. The authors of the systematic review did not report individual characteristics of the studies. The sensitivity and specificity of <sup>18</sup>F-FDG PET were greater than 85% and 67%, respectively. The sensitivity was similar to that of CT or MRI ( $P = 0.46$ ), but the specificity was higher in the PET group ( $P = 0.06$ ).

The panel concluded that, overall, the data are too uncertain to support the use of PET in determining the anatomic extent of head and neck cancer. However, none of the studies used PET as the initial approach, reserving endoscopy for patients with PET-negative findings.

*Areas of Uncertainty and Implications for Future Research.* Despite consistency in the available evidence, further research is likely to alter confidence in the effect of <sup>18</sup>F-FDG PET for this use and may change the estimate. A study evaluating negative PET findings followed by surgical correlation will be the most useful.

### **Is <sup>18</sup>F-FDG PET Useful for Staging Head and Neck Cancer?**

*Background and Rationale.* The standard imaging tests for staging head and neck tumors are CT and MRI. These methods often results in under- or overstaging. Better imaging methods will be useful.

*Evidence Summary.* The panel concluded that PET should routinely be added to CT or MRI to improve nodal or distant-disease staging of head and neck cancer. On the basis of 3 systematic reviews of the evidence, the panel judged the evidence to be sufficient (moderate-quality) that the addition of PET to CT or MRI will likely improve important health-care outcomes and concluded that <sup>18</sup>F-FDG PET is beneficial.

*Status of the Evidence.* The panel identified 3 systematic reviews (55,57,63). The overall quality of the evidence was moderate. Two systematic reviews analyzed detection of regional local metastases and compared PET with CT or MRI. Both systematic reviews included 17 studies. The panel could not exclude the possibility of overlap between the studies. The quality of the systematic reviews themselves was low for one (57) and high for the other (55). The quality of primary research evidence was moderate. No

randomized studies were performed. One of the systematic reviews analyzed 540 patients, and the other did not state the number of patients. In general, the studies in the 2 reviews suffered from spectrum, verification, and detection biases. The sensitivity and specificity of  $^{18}\text{F}$ -FDG PET were greater than those of CT and MRI. One of the reviews stated that 6 studies evaluated disagreements between PET and other imaging tests and concluded that PET was usually correct among the discordant findings, in 60%–100% of the cases (55). One review (63) evaluated the overall detection of metastases (local and distant) with PET; this review included 7 prospective studies, which analyzed 30–78 patients per study. The quality of the systematic review was moderate; the quality of primary research evidence was also moderate. Of the 7 studies, 4 compared PET with CT or MRI and found a PET sensitivity of 72%–87% and specificity of 92%–100%. PET was superior to CT in terms of PPV and NPV in 2 studies (PET PPV, 90% and 89%; PET NPV, 93% and 99%; CT PPV, 40% and 74%; CT NPV, 72% and 95%) (63). An additional systematic review (57) evaluated the use of PET in the detection of distant metastases and synchronous primaries in patients diagnosed with primary squamous cell cancer of the head and neck. The quality of the systematic review and the primary evidence was low. Potential bias was found in this systematic review because of mixed analysis of patient and lesion data and the fact that the reference test was not clearly stated. This systematic review included 4 studies, which analyzed 12–59 patients per study. In the largest study, which was retrospective, PET was more accurate than bronchoscopy for the detection of synchronous lung lesions (80% vs. 50%).

The panel concluded that the use of PET in addition to conventional imaging tests is beneficial for local staging of head and neck cancer. No systematic review reported subgroup analysis according to lymph node status (i.e., according to clinical N0 vs. non-N0 neck lymph node involvement). The indirect evidence of benefits stems from the detection of locoregional metastases, which, in turn, should direct physicians to the proper treatment. PET can help some patients avoid cervical lymph node dissection (which can be disfiguring).

However, despite the potential benefit in detecting lesions below the neck, the net benefit of using PET for detecting distant metastases is still uncertain. Nevertheless, the panel believed that the use of PET might be beneficial in patients with advanced-stage disease, in whom the odds of having a distant metastasis are greater. However, in these cases, PET-positive lesions should undergo biopsy.

*Areas of Uncertainty and Implications for Future Research.* Despite the consistency in the available evidence, further research is likely to alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use and may change the estimate. In particular, better studies are needed of the impact on patient management of detecting metastases below the neck.

## Is $^{18}\text{F}$ -FDG PET Useful for Detecting Recurrence of Head and Neck Cancer?

*Background and Rationale.* The standard imaging tests for recurrence of head and neck tumors are CT and MRI. These methods often give false-positive or false-negative results. Better imaging methods would permit more appropriate treatment.

*Evidence Summary.* The panel concluded that PET should routinely be added to conventional imaging in the diagnostic work-up of patients with a potential recurrence of head and neck cancer. The panel found moderate evidence that PET will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial.

*Status of the Evidence.* The panel identified 3 systematic reviews (38,57,63) that evaluated the presence of recurrent or residual disease. The first systematic review analyzed 15 studies (57), which compared PET with CT or MRI. The authors found a PET sensitivity of 73%–100% and a CT or MRI sensitivity of 25%–100% ( $P = 0.01$ ). The specificities were 57%–100% for PET, versus 33%–100% for CT or MRI ( $P = 0.02$ ).

The second systematic review (63) found 2 prospective studies. The first study compared PET, CT, and clinical examination in stages III and IV head and neck cancer. In that study, which identified recurrence at 1 y, sensitivity was 100% for PET, compared with 38% for CT and 44% for clinical examination. The authors concluded that specificity was good or excellent using all 3 methods but did not provide data. The second prospective study enrolled 44 patients and demonstrated a sensitivity of 96% for PET, versus 73% for CT or MRI. PET had a specificity of 61%, versus 50% for CT or MRI.

Finally, Facey et al. (38) identified 3 scenarios in which  $^{18}\text{F}$ -FDG PET was used to detect recurrence. However, in all these scenarios, the quality of the primary studies was unclear or low, and whether the studies of this or the other 2 systematic reviews overlapped was not clear.

For the first of these scenarios—restaging during follow-up after primary treatment with radiation or surgery for head and neck cancer—24 studies were found and 18 studies had patients as the unit of analysis, enrolling 10–48 patients per study. The pooled sensitivity was 90% (range, 33%–100%) and specificity was 76% (95% CI not reported). PET was considered better than the comparators in 6 studies ( $n = 140$ ), had mixed or neutral results in 4 studies ( $n = 152$ ), and was worse than CT in 1 study ( $n = 13$ ). The BCBSA also examined the same studies and reported the same findings in this setting (55). Only 1 study specifically addressed a change in management that was due to PET findings. In that study, PET had been used to support palliative care instead of curative surgery in 9 of the 29 patients (55).

For the second scenario—assessment of residual or recurrent head and neck cancer—the authors identified 15 studies comparing PET with CT or MRI and enrolling 10–66 patients per study. PET had a sensitivity better than 85% in 14 of 15 studies and a specificity better than 80% in 10 of

15 studies. Three studies carefully described change-in-management decisions based on PET findings. In the first study, PET correctly indicated the need for biopsy in 16 of 17 patients, versus 11 of 17 patients for CT or MRI. PET helped avoid biopsy in 14 of 21 patients. In the second study, distant metastases were identified by PET in 7 of 22 patients and resulted in a change in management (from surgery to palliative treatment). The third study described a decision to change the management for 26 of 66 patients after PET, and 23 of these decisions were found to be correct.

For the third scenario—restaging regional lymph nodes in patients with recurrent head and neck cancer (investigation at follow-up visit)—the authors described 10 studies enrolling a total of 350 patients (13–50 patients per study). PET had a sensitivity of 88% and a specificity of 78%, the reported LR+ was 4.0 (95% CI, 2.8–5.6), and the reported LR– was 0.16 (95% CI, 0.10–0.25).

The overall quality of the evidence was moderate, the quality of the systematic reviews themselves was either unclear or moderate, and the quality of primary research evidence was moderate. Major deficiencies were small sample sizes, a lack of standardization between studies, and the possibility of spectrum, verification, and detection biases. In all studies, sensitivity and specificity were higher for PET than for other imaging tests.

The panel concluded that adding PET to conventional imaging tests is beneficial for detecting recurrence of head and neck cancer. The greatest benefit was attributed to accurate detection of cancer. This accuracy, in turn, should help physicians choose the appropriate treatment.

*Areas of Uncertainty and Implications for Future Research.* Despite consistency in the available evidence, further research is likely to alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use and may change the estimate. In particular, high-quality studies evaluating the impact of  $^{18}\text{F}$ -FDG PET on treatment selection and, consequently, on patient outcomes will be useful.

## LUNG CANCER

### Is $^{18}\text{F}$ -FDG PET Useful for Differentiating Benign from Malignant Lesions, Including Evaluation of SPN?

*Background and Rationale.* An SPN, or “coin lesion,” is an approximately round lesion that is less than 3 cm in diameter and completely surrounded by pulmonary parenchyma, without other abnormalities. Lesions larger than 3 cm are called masses and are often malignant. The incidence of cancer in patients with solitary nodules is 10%–70% (64). In fact, Henschke et al. saw at least 1 nodule on baseline chest radiography in almost 7% of healthy volunteers (who were primarily older smokers or former smokers) in a lung cancer screening trial (65).  $^{18}\text{F}$ -FDG PET might have an important role in the differential diagnosis of SPN.

*Evidence Summary.* The panel concluded that  $^{18}\text{F}$ -FDG PET should routinely be obtained in the diagnostic work-up

of patients with SPN. The panel found moderate-quality evidence that the use of  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries in low-risk patients and enabling curative surgeries in high-risk patients. In low-risk patients, the NPV of  $^{18}\text{F}$ -FDG PET is extremely high; therefore, negative  $^{18}\text{F}$ -FDG PET findings can be followed up with observation only. However, in high-risk patients, even negative findings should be followed by histopathologic investigation. The panel does not endorse specific quantitative thresholds above or below which a patient should be considered at a high or low risk, respectively, for lung cancer. The judgments rendered were qualitative.

*Status of the Evidence.* The panel identified 3 systematic reviews (63,66,67), which analyzed 4–46 studies. The overall quality of the evidence was moderate, the quality of the systematic reviews themselves ranged from unclear to high, and the quality of primary research evidence was moderate. Gould et al. (67) included 40 studies in their review, 20 of which were prospective. This systematic review (67) used a quality index score to assess the primary research studies and verified that 14 studies satisfied 70%–80% of the quality criteria, 18 satisfied 50%–69%, and 5 satisfied less than 50%. The major deficiencies were a failure to adequately describe the PET test procedure (including a reference to glucose level and how the PET images were read) and the reference test. Most studies enrolled patients with pulmonary nodules or masses, but 6 enrolled only patients with nodules and 7 provided separate results for nodules. Overall, the median prevalence of malignancy was 72.5%. The summary log odds ratio for PET was 4.68 (95% CI, 4.21–5.14), corresponding to a maximum joint sensitivity and specificity of 91.2% (95% CI, 89.1%–92.9%).

For focal pulmonary lesions of any size ( $n = 1,474$ ), PET sensitivity was 83%–100% (mean, 96%). The specificity was extremely variable, at 0%–100% (mean, 73%). When analyzed by pulmonary nodules only ( $n = 450$ ), the mean sensitivity and specificity were 94% and 86%, respectively, and the median sensitivity and specificity were 98% and 83%, respectively. The summary log odds ratio for  $^{18}\text{F}$ -FDG PET was 4.40 (95% CI, 3.70–5.09), corresponding to a maximum joint sensitivity and specificity of 90.0% (95% CI, 86.4%–92.7%). The authors of this review found no difference between the accuracy of  $^{18}\text{F}$ -FDG PET for pulmonary nodules and the accuracy for pulmonary lesions of any size ( $P = 0.43$ ). However, only 8 studies evaluated lesions smaller than 1 cm. For low-risk patients, the posttest likelihood of malignancy with a negative PET finding was 1%. Therefore, the authors concluded that the NPV of PET in such a population would be sufficient to recommend observation and follow-up. However, high-risk patients had a lower NPV (86%) with negative PET findings. Therefore, further diagnostic investigation is indicated.

In their review, Fischer et al. (66) analyzed 800 patients in diagnostic studies and found a sensitivity of 96% and a specificity of 78% for  $^{18}\text{F}$ -FDG PET. The risk of false-positive

findings was high. The PPV was 91% and the NPV was 90%. The LR+ was 4.4 and the LR- was 0.05.

The third systematic review (63) included 4 prospective studies assessing the effectiveness of PET in differentiating malignant from benign lesions when CT-guided biopsy had failed to provide a final diagnosis or when the procedure was contraindicated. The sample size from the included studies was 50–109 patients per study. Sensitivity was 86%–100%, specificity was 40%–90%, PPV was 88%–95%, and NPV was 55%–100%. Several commentators emphasized that the predictive values of PET depend on the prior probability of lung cancer in the population of interest. PPV will be higher in high-risk patients, whereas NPV will be higher in low-risk patients. Thus, PET will be most useful in patients at an intermediate risk of lung cancer. Quantitative models integrating age, smoking history, size of tumors, and other variables were proposed to estimate the likelihood of lung cancer in a particular patient (68,69). The precise cutoff points between which PET should be ordered, and below which lung cancer can be considered ruled out or above which ruled in, were also proposed on the basis of the cost-effectiveness analysis (70). Physicians should use their judgment when determining patient risk. Detterbeck et al. (71), in their narrative review, indicated that PET is associated with high false-positive rates in patients clinically suspected of infection, whereas false-negative rates are high in patients with bronchioloalveolar carcinoma or carcinoid tumors. These considerations are likely to be helpful to physicians as they decide whether to administer PET to patients with SPN. When using PET, individual physicians should still apply sound judgment according to the general principles of ordering and interpreting diagnostic tests.

The panel concluded that, overall, using PET in the diagnostic work-up of patients with SPN is beneficial. The greatest benefit was seen in the early detection of potential malignant lesions or in the elimination of suggestive lesions in low-risk patients. These benefits, in turn, should help avoid futile surgeries in low-risk patients and enable curative surgeries in high-risk patients.

*Areas of Uncertainty and Implications for Future Research.* Because the available evidence was consistent, further research probably will not alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use and may not change the estimate. However, further research integrating quantitative decision-support systems to help tailor the use of PET to each patient's circumstance needs to continue and will likely be useful. In addition, the role of integrated PET/CT has not been systematically evaluated in this setting, and this technology may further improve discernment of benign from malignant lesions in high-risk patients.

### **Is $^{18}\text{F}$ -FDG PET Useful for Staging Non-Small Cell Lung Cancer (NSCLC)?**

*Background and Rationale.* Lung carcinoma is the leading cause of cancer-related death in the Western world (41). Accurate staging is essential because distant metastases and

metastases to mediastinal lymph nodes have a crucial impact on the prognosis of non-small-cell lung cancer, making accurate staging fundamental for selecting the best therapy (72).

*Evidence Summary.* The panel concluded that PET should routinely be added to the conventional work-up of NSCLC patients. The panel found high-quality evidence that PET will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries through detection of extrathoracic disease. Evidence from RCTs shows that PET alone (without a conventional work-up) is not better than a conventional work-up alone, and therefore PET is not recommended as a single imaging modality.

*Status of the Evidence.* The panel identified 5 systematic reviews (30,66,73–75), which analyzed 14–46 studies. The overall quality of the evidence was high, the quality of the systematic reviews themselves ranged from unclear to high, and the quality of primary research evidence ranged from low to high. The latest review was conducted by Birim et al. (73) and included studies until 2003. That review used a quality index score to assess the primary research studies. Only 4 studies met all the quality criteria. The item most often poorly described was the study population. However, results were consistent regardless of the quality of the original studies and subgroup analyses. The systematic review by Birim et al. also suffered from potential biases, including publication and verification biases because the analysis mainly included operable patients. The review compared PET and CT in detecting mediastinal metastases of NSCLC. The authors identified 17 eligible studies (15 prospective studies) totaling 833 patients. Five studies evaluated pulmonary nodules, lymph node metastases, and distant metastases; 3 evaluated pulmonary nodules and mediastinal metastases; 3 evaluated lymph node metastases; and 6 evaluated only mediastinal lymph nodes. For detection of mediastinal lymph node metastases, PET had an overall sensitivity of 83% (95% CI, 77%–87%) and an overall specificity of 92% (95% CI, 89%–95%). On the other hand, for detection of mediastinal lymph node metastases, CT had an overall sensitivity of 59% (95% CI, 50%–67%) and an overall specificity of 78% (95% CI, 70%–84%). No statistically significant heterogeneity in sensitivity or specificity was detected between the 2 methods. For detection of mediastinal lymph nodes, the PET ROC was 0.90 (95% CI, 0.86–0.95) and the CT ROC was 70% (95% CI, 0.65–0.75) ( $P < 0.0001$ ).

Gould et al. (30) also studied the usefulness of  $^{18}\text{F}$ -FDG PET in mediastinal staging and concluded that PET was more sensitive but less specific when CT showed lymph node enlargement than when it did not ( $P < 0.002$ ).

The panel also identified 3 RCTs that addressed the use of  $^{18}\text{F}$ -FDG PET in the staging of lung cancer. All 3 RCTs were of high quality and provided direct evidence on patient outcomes.

The PLUS trial (76) studied the effect of PET in the reduction of futile thoracotomies in patients with suspected



NSCLC who were scheduled for surgery after the conventional work-up. This trial allocated 188 patients (few patients had metastatic disease), 96 individuals in the conventional work-up arm and 92 in the conventional work-up + PET arm. The authors defined futile surgeries as surgeries that were performed for benign lesions; surgeries that were performed when there was histopathologically proven mediastinal lymph node involvement (stage IIIA-N2), or when stage IIIB was present; exploratory thoracotomies performed for any other reason; surgeries performed in cases of recurrent disease; or surgeries performed when death from any cause occurred within 1 y of randomization. This trial showed a significant number of patients with futile surgery in the conventional work-up arm, versus the number in the conventional work-up + PET arm (relative risk reduction, 51% [95% CI, 32%–80%] [ $P = 0.003$ ] in favor of PET). In the conventional work-up arm, 39 patients had futile surgeries, and in the conventional work-up + PET arm, only 19 patients had futile surgeries. Nineteen recurrences or deaths occurred within 1 y of futile surgery in the conventional work-up group, versus 10 in the conventional work-up + PET group.

Viney et al. (77) addressed the impact of  $^{18}\text{F}$ -FDG PET on the clinical management and surgical outcome of patients with stage I–II NSCLC. The authors wanted to verify that the addition of  $^{18}\text{F}$ -FDG PET would reduce the number of unnecessary thoracotomies in those patients. They included 183 patients, 92 in the conventional work-up arm and 91 in the conventional work-up + PET arm.  $^{18}\text{F}$ -FDG PET resulted in further investigation or other changes in the management of 12 patients (13%) ( $P = 0.2$ ).  $^{18}\text{F}$ -FDG PET could have a potential impact on management in 26% of patients. The sensitivity and specificity of  $^{18}\text{F}$ -FDG PET for the detection of mediastinal disease were 73% and 90%, respectively. With a minimum 1-y survival, 80% of patients were alive in the PET arm and 77% in the no-PET arm ( $P$  not shown).

The third RCT, conducted by Herder et al. (78,79), compared  $^{18}\text{F}$ -FDG PET alone with conventional work-up for the diagnosis and staging of NSCLC. They included 465 patients, 232 in the PET arm and 233 in the CT arm. The proportion of patients requiring at least 3 tests was 51% in the conventional work-up group, compared with 51% in the PET group ( $P = 0.82$ ). The number of thoracotomies was also similar in the 2 groups (41% for PET, vs. 38% for CT). The requirement for one or more invasive tests for N staging was inferior in the  $^{18}\text{F}$ -FDG PET group (22% for PET, vs. 39% for CT,  $P = 0.0001$ ). Also, differences in the costs of diagnostic procedures between the 2 arms were not statistically significant. The authors concluded that the application of PET upfront in the staging of patients with (suspected) NSCLC carries an overall accuracy similar to that of a conventional work-up, with a limited impact on test substitutions. Consequently, this finding reinforces the suggestion that  $^{18}\text{F}$ -FDG PET should be used as an add-on tool for assessing lung cancer and should not replace the conventional work-up for the diagnosis and staging of lung cancer.

The panel concluded that adding PET to CT is beneficial in the mediastinal diagnostic work-up of patients with lung cancer. Indirect evidence indicated that the greatest benefit was due to metastasis detection, which generally precluded surgery and should, in turn, help avoid futile surgeries and help physicians choose the appropriate treatment. Direct evidence from randomized trials further showed that adding  $^{18}\text{F}$ -FDG PET to the conventional work-up can decrease futile surgeries and has a positive impact on the management of lung cancer patients.

*Areas of Uncertainty and Implications for Future Research.* The evidence about the use of PET in this setting is reliable and consistent. Therefore, further research most likely will not alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use and may not change the estimate. Nevertheless, the role of integrated PET/CT has not been systematically evaluated for this indication, and this technique might further improve confidence in the use of  $^{18}\text{F}$ -FDG PET for staging lung cancer.

#### **Is $^{18}\text{F}$ -FDG PET Useful for Detecting Distant Metastases in Patients with Proven or Suspected NSCLC?**

*Evidence Summary.* The panel concluded that  $^{18}\text{F}$ -FDG PET should be obtained in the diagnostic work-up of lung cancer patients for distant metastases. The panel found moderate-quality evidence that  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries.

*Status of the Evidence.* The panel identified 1 systematic review (38), which analyzed 19 studies enrolling 21–167 patients per study. The overall quality of the evidence was moderate, the quality of the systematic review itself was unclear, and the quality of primary research evidence was moderate. The major deficiency was the possibility of spectrum bias. The systematic review found that  $^{18}\text{F}$ -FDG PET detected 10%–20% more distant metastases than did other imaging methods. Also, 16 of the studies evaluated change-in-management outcome. They showed a management change in 9%–64% of patients and that, in most cases, the patients were not taken to surgery.

The panel concluded that, overall, using PET in the metastatic work-up is beneficial. The greatest benefit was attributed to detection of these distant metastases. This detection, in turn, should help avoid futile surgeries and help physicians administer the appropriate treatment.

*Areas of Uncertainty and Implications for Future Research.* The available evidence is consistent. Therefore, further research most likely will not alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use and may not change the estimate. Nevertheless, the role of integrated PET/CT has not been systematically evaluated for this indication, and this technique may further improve confidence in the use of  $^{18}\text{F}$ -FDG PET for staging lung cancer.

### Is <sup>18</sup>F-FDG PET Useful for Diagnosing and Managing Small Cell Lung Cancer (SCLC)?

*Evidence Summary.* The panel concluded that the evidence is insufficient, of too poor a quality, or too inconsistent to support the use of <sup>18</sup>F-FDG PET in the management of SCLC. <sup>18</sup>F-FDG PET may eventually have a role in the evaluation of limited disease, but more solid evidence is required before such a recommendation can be made.

*Status of the Evidence.* The panel identified 1 systematic review (38), which addressed 3 aspects of the use of <sup>18</sup>F-FDG PET in SCLC. The overall quality of the evidence was low in all 3 scenarios, the quality of the systematic review itself was unclear, and the quality of primary research evidence was low. No randomized studies were performed, and the review did not make clear how many studies were retrospective and how many were prospective and enrolling consecutive patients.

The first scenario addressed the use of <sup>18</sup>F-FDG PET for the diagnosis of occult SCLC in patients with suspected paraneuroplastic neurologic syndrome in whom conventional imaging findings were negative. Facey et al. (38) found only 1 study enrolling 43 patients, with available data on 39 patients. Only 5 patients had the condition of interest, and therefore these findings represent very preliminary data.

The second scenario addressed the use of <sup>18</sup>F-FDG PET for staging SCLC through determining the extent of disease. Five studies were found, enrolling 3–30 patients per study. The sensitivity was 89%–100% and the specificity was 100%. In the largest study (30 patients), the CT or MRI comparator had a sensitivity of 65% and a specificity of 100%. In the second largest study (25 patients), the CT comparator had a sensitivity of 93% and a specificity of 90%.

The final scenario was related to the use of <sup>18</sup>F-FDG PET for restaging after initial chemotherapy or radiation treatment of SCLC, to detect residual disease or a new site. Only 2 small studies were found, and each reported different outcomes. Recurrence was investigated in a study that included only 12 patients; PET had a sensitivity of 100% and a specificity of 80% for the detection of recurrence.

*Areas of Uncertainty and Implications for Future Research.* Because of the lack of reliable evidence, further research is likely to alter confidence in the effect of <sup>18</sup>F-FDG PET for this use and may change the estimate. High-quality studies are needed in this setting. Similarly, the role of integrated PET/CT has not been systematically evaluated for this indication.

## LYMPHOMA

### Is <sup>18</sup>F-FDG PET Useful for Staging Lymphoma?

*Background and Rationale.* Staging is extremely relevant to the treatment of any type of cancer but is critically important for patients with lymphoma. Proper staging allows individualization of specific treatments to match the extent of disease. This specificity, in turn, may help avoid administration of unnecessary toxic treatments such as extended-field radiation or overly aggressive chemotherapy, further result-

ing in a decrease in the risk of secondary malignancies and other sequelae (80). <sup>18</sup>F-FDG PET can help better estimate the extent of disease in lymphoma patients, consequently leading to better management and outcomes.

*Evidence Summary.* The panel suggested that <sup>18</sup>F-FDG PET should routinely be obtained in addition to the conventional work-up in the pretreatment staging of lymphoma. The panel found low-quality but relatively consistent evidence that the addition of <sup>18</sup>F-FDG PET will improve important health-care outcomes and judged that <sup>18</sup>F-FDG PET is probably beneficial in this setting. <sup>18</sup>F-FDG PET is considered more valuable in Hodgkin's disease (HD) and early-stage aggressive non-Hodgkin's lymphoma (NHL) and less useful in indolent NHL. Therefore, depending on clinical circumstances, physicians may decide to modify this recommendation.

*Status of the Evidence.* The panel identified 5 systematic reviews (38,63,81–83), which analyzed 4–20 studies. The overall quality of the evidence was low, the quality of the systematic reviews themselves varied from low to high, and the quality of primary research evidence varied from low to moderate. No randomized studies were performed. The reviews did not make clear how many studies were retrospective and how many were prospective and enrolling consecutive patients. The major deficiencies were verification, timing, detection, and spectrum biases and deficiencies in describing the impact of the PET findings on patient management. In addition, almost all studies enrolled patients with various histologic types, raising the possibility of significant clinical heterogeneity among studies and results. The heterogeneity of reported evidence in lymphoma is a major problem, making it all but impossible to adequately synthesize the research.

The BCBSA review (83) demonstrated that PET had better overall diagnostic accuracy than did CT in all studies that assessed both techniques. Eleven studies evaluated alterations in patient management, and 5 of these 11 reported change-in-management information. <sup>18</sup>F-FDG PET resulted in a change in management in 8%–20% of patients. Ten studies reported concordance between the PET results and other imaging results. <sup>18</sup>F-FDG PET was discordant with the conventional work-up in 11%–55% of patients, and among the discordances, PET was accurate in 40%–96% of cases.

Hutchings et al. (81) also studied the role of <sup>18</sup>F-FDG PET in staging lymphoma and identified 6 studies that analyzed a mixed patient population. Of the 6 studies, 2 clearly reported a change in management corresponding to 8% in nodal disease and 16% in extranodal disease. Despite technical differences between the <sup>18</sup>F-FDG PET protocols, PET had a higher diagnostic sensitivity than did conventional staging procedures. The same authors also reported the <sup>18</sup>F-FDG PET results exclusively in an HD population. They identified 7 studies and found a change in management in 3%–25% of the patients.

Facey et al. (38) addressed the impact of <sup>18</sup>F-FDG PET in the identification of more advanced, nonbulky or bulky

disease to determine the initial therapy. They identified 7 studies enrolling 11–93 patients per study. Two studies (with 52 and 76 patients) compared PET with biopsy or scintigraphy. All had a specificity greater than 90% and a sensitivity of 79%–100%. One study (93 patients) used  $^{67}\text{Ga}$  scintigraphy as a comparator; sensitivity was greater than 85% for PET and the comparator, and specificity was not reported. Only 2 small studies used CT as a comparator (27 patients total). Eleven reports indicated how PET changed the staging, and some indicated how this change affected management. Two studies were well reported. The first study, enrolling 50 patients, compared PET with gallium scintigraphy, resulting in upstaging using PET in 8 cases and using gallium scintigraphy in 7 cases. The change-in-management outcomes were recorded in 10 PET cases and in 7 gallium scintigraphy cases. The second study, enrolling 49 patients, found upstaging in 27 PET cases and downstaging in 2 PET cases. Interestingly, all but 1 patient were treated according to PET staging. Many studies could also be duplicated in other reviews. Evidence about change in management was typically related to few patients in each study, with few details given.

Finally, the most recent systematic review (82) evaluated the effectiveness of PET in staging and restaging lymphoma. This high-quality systematic review evaluated 20 studies (7 prospective). Patient data and lesion data were extracted from primary studies. For the studies that reported patient-based data, the pooled sensitivity was 91% (95% CI, 88%–93%) and the pooled false-positive rate was 10% (95% CI, 7%–14%). The maximum joint sensitivity and specificity was 88% (95% CI, 85.0%–91%). The pooled sensitivity and false-positive rate appeared to be higher in patients with HD than in those with NHL; however, the number of studies that exclusively evaluated NHL was limited to 3. The overall change-in-management rate was 30% as a result of PET findings. This systematic review confirmed the problems with the quality of primary research evidence but concluded that PET is useful as an add-on test in staging lymphoma.

The panel agreed with the authors of all reviews that more and better evidence is needed. CT may be superior for the detection of intraabdominal lymph nodes. However,  $^{18}\text{F}$ -FDG PET appears to have a higher sensitivity for the detection of nodal and extranodal disease. PET also had a better PPV than did bone scanning in identifying bone involvement. Because of false-negative results, PET should be considered an additional imaging tool and not be used alone. Because of false-positive results, PET cannot replace biopsy.  $^{18}\text{F}$ -FDG PET is also not considered reliable for the staging of low-grade NHL.

The panel concluded that, overall, using  $^{18}\text{F}$ -FDG PET in the staging of lymphoma patients is beneficial. The greatest benefit is attributed to more accurate detection of the extent of disease than is possible with conventional imaging alone. This accuracy is expected to enable administration of treatments appropriate to the level of disease, ultimately improving patient outcomes.

*Areas of Uncertainty and Implications for Future Research.* Further research is likely to alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use and may change the estimate. Further studies comparing conventional imaging with added PET should be performed both for restaging and for monitoring treatment response. Integrated PET/CT scanners may overcome some of the problems with PET alone.

### **Is PET Useful for Evaluating Bone Marrow Infiltration in the Staging of Lymphoma?**

*Background and Rationale.* The assessment of bone marrow infiltration is an integral part of staging lymphoma. The usual standard procedure for detection of bone marrow lesions is bone marrow biopsy. However, this procedure has several disadvantages (e.g., it is a painful procedure with potential complications, and it can miss the lesion of interest and thus is not suitable for evaluating multiple-site involvement). In recent years, the use of  $^{18}\text{F}$ -FDG PET for staging bone marrow involvement in lymphoma has been studied.

*Evidence Summary.* The panel concluded that  $^{18}\text{F}$ -FDG PET may be added to bone marrow biopsy for staging and restaging lymphoma. The panel found moderate-quality evidence that adding  $^{18}\text{F}$ -FDG PET will improve important health-care outcomes and judged that  $^{18}\text{F}$ -FDG PET is probably beneficial in this setting. Biopsy should be directed to PET-positive sites (if the site is deemed easily accessible). Because of the high proportion of false-negative results, conventional masked biopsies are still needed.

*Status of the Evidence.* The panel identified 1 systematic review (84), which analyzed 13 studies enrolling 7–105 patients per study. The overall quality of the evidence was moderate, the quality of the systematic review itself was high, and the quality of primary research evidence varied from low to moderate. No randomized studies were performed. Seven studies enrolled patients prospectively. Eight studies described masked interpretation of PET images. At least half the studies included mixed populations (primary vs. recurrent lymphomas and HD vs. NHL). The  $^{18}\text{F}$ -FDG PET sensitivity was 0%–100%, and the specificity was 72%–100%. The overall sensitivity and specificity were 51% (95% CI, 38%–64%) and 91% (95% CI, 85%–95%), respectively. The LR+ was 5.75 (95% CI, 3.85–9.48), and the LR– was 0.67 (95% CI, 0.55–0.82). Only half the patients in whom bone marrow infiltration was detected on bone marrow biopsy had tumor detected on PET. More than 90% of patients with negative findings on bone marrow biopsy also had negative findings on  $^{18}\text{F}$ -FDG PET.

The panel concluded that using  $^{18}\text{F}$ -FDG PET for assessing bone marrow involvement during the staging of lymphoma is beneficial. The greatest benefit was attributed to improved staging because of the weakness of the current reference standard. Therefore, the results of  $^{18}\text{F}$ -FDG PET may enable treatment appropriate to the level of disease, ultimately leading to improved patient outcomes.

*Areas of Uncertainty and Implications for Future Research.* Further research is likely to alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use. Prospective studies enrolling a more homogeneous disease-specific population (e.g., HD, low-grade, or high-grade lymphoma) are needed.

### **Is $^{18}\text{F}$ -FDG PET Useful for Restaging or Detecting Relapse, Assessing Residual Mass, or Assessing Progression After Completion of Initial Treatment in Lymphoma Patients?**

*Background and Rationale.* Many lymphoma patients present with residual masses after completing induction therapy, but less than 20% of them will eventually experience relapse (85). These residual masses may consist of necrotic tissue or active disease, and conventional imaging tests, such as CT or MRI, may not be able to detect the difference. Therefore,  $^{18}\text{F}$ -FDG PET might be helpful if it is more accurate than conventional radiologic imaging techniques for restaging after completion of initial treatment, ultimately enabling administration of treatment according to the actual status of the disease (85).

*Evidence Summary.* For HD, the panel concluded that PET should routinely be added to the conventional work-up for restaging or detecting recurrence in patients to whom curative treatment was administered. However, if the  $^{18}\text{F}$ -FDG PET findings are positive, further confirmation by biopsy is mandatory. The panel found moderate evidence that adding  $^{18}\text{F}$ -FDG PET in this clinical setting will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial.

For NHL, the panel concluded that PET should routinely be added to the conventional work-up for restaging or detecting recurrence in patients who were treated with curative intent. However, further confirmation by biopsy, if  $^{18}\text{F}$ -FDG PET findings are positive, is mandatory. The panel found low but consistent evidence that adding  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial. Depending on clinical circumstances, physicians may decide to modify this recommendation. For example, if palliative management is the goal,  $^{18}\text{F}$ -FDG PET is not indicated. Likewise, PET is not indicated in indolent NHL.

*Status of the Evidence.* The panel identified 2 systematic reviews (38,81), which analyzed 1–8 studies. The overall quality of the evidence was low for NHL and moderate for HD, the quality of the systematic reviews themselves was high or unclear, and the quality of primary research evidence varied from unclear to moderate. No randomized studies were performed on the topic.

Facey et al. (38) addressed the impact that restaging lymphoma had on identifying residual tumor masses, after a partial or complete response to induction therapy, to avoid unnecessary consolidation radiotherapy if no active residual disease was present. The authors identified 8 PET studies and 6 CT studies. In patients with positive CT findings (a total of

246 patients from 7 studies),  $^{18}\text{F}$ -FDG PET had a sensitivity of 80% (95% CI, 59%–94%) and a specificity of 89% (95% CI, 74%–97%). In patients without CT (a total of 384 patients from 7 studies), the  $^{18}\text{F}$ -FDG PET sensitivity was 81% (95% CI, 63%–92%) and specificity was 95% (95% CI, 90%–99%). The overall CT sensitivity was 75% (95% CI, 58%–88%) and specificity was 45% (95% CI, 27%–64%), based on data from 6 studies enrolling 266 patients. The data appeared to indicate that any HD patient with a residual mass and negative PET findings is unlikely to experience relapse. PET, however, has a high false-positive rate, and biopsy is mandatory before treatment is reinitiated.

The panel concluded that using  $^{18}\text{F}$ -FDG PET is beneficial in restaging or detecting progression of lymphoma (after completion of initial treatment). The greatest benefit was attributed to more accurate detection of the extent of disease at staging or restaging, including better differentiation of necrotic or scar tissue from active disease in patients with a residual mass. Therefore, the results of  $^{18}\text{F}$ -FDG PET could enable the administration of treatment appropriate to the level of disease, ultimately leading to improved patient outcomes.

*Areas of Uncertainty and Implications for Future Research.* Further research is likely to alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use. The optimal timing of PET has not been established. Further studies comparing conventional imaging with added PET should be performed both for restaging and for monitoring treatment response.

### **Is $^{18}\text{F}$ -FDG PET Useful for Following Up and Diagnosing Relapse in Lymphoma Patients?**

*Background and Rationale.* Detection of early, preclinical relapse would be useful if appropriate salvage therapy could be administered earlier, ultimately improving survival. This advantage could be particularly important in HD (81). Currently, no evidence-based data support the use of  $^{18}\text{F}$ -FDG PET for following up lymphoma patients after successful first-line therapy (81). The use of  $^{18}\text{F}$ -FDG PET could potentially be beneficial in this setting.

*Evidence Summary.* The panel concluded against routine administration of PET for detecting relapse in asymptomatic HD or NHL and found limited evidence supporting the use of  $^{18}\text{F}$ -FDG PET in the routine follow-up of asymptomatic patients. A negative result would not affect the follow-up strategy, and positive results could easily be false-positive.

*Status of the Evidence.* For HD, data are sparse on the utility of  $^{18}\text{F}$ -FDG PET in following up and diagnosing relapse. Hutchings et al. (81) found only 1 study, which enrolled 36 consecutive patients who underwent  $^{18}\text{F}$ -FDG PET 1 mo after the end of treatment and then every 4–6 mo for 2–3 y. Conventional work-up identified a residual mass in 19 patients, 14 of whom had negative PET findings and never experienced relapse. In total, 11 patients had positive PET findings, with only 5 experiencing relapse, thus indicating a 55% (6/11) false-positive relapse rate. Only 2 of 5 patients had clinical symptoms at the time of relapse.

For NHL, the panel identified no systematic review addressing this question. A narrative review by Burton et al. (86) indicated that only limited evidence is available about the value of PET in following up NHL patients.

*Areas of Uncertainty and Implications for Future Research.* Further research is likely to alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use and will likely change the current estimate. Most informative would be a trial comparing the impact of administering early salvage therapy based on  $^{18}\text{F}$ -FDG PET findings (with or without conventional imaging studies) versus therapy when a patient clinically relapses.

## MELANOMA

### Is $^{18}\text{F}$ -FDG PET Useful for Detecting Metastases of Melanoma?

*Background and Rationale.* Cutaneous melanoma ranks fifth in cancer incidence among men and seventh among women in the United States (87). It frequently metastasizes and is difficult to treat. Accurate staging is important for optimizing therapy and selecting appropriate patients for experimental trials.  $^{18}\text{F}$ -FDG PET has been studied extensively in the last 2 decades as a potential tool to help in the detection of metastatic cutaneous melanoma. Some have also argued that PET could be useful in patients who are at a high risk for systemic relapse and are being considered for aggressive medical therapy (88) and as an additional imaging tool for detecting recurrence.

*Evidence Summary.* The panel concluded that  $^{18}\text{F}$ -FDG PET should routinely be added to conventional imaging for staging and detecting recurrent melanoma. The panel found moderate evidence that  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial, mostly by helping tailor treatment toward the stage of disease. However, physicians should be aware that pulmonary or brain metastases might not be accurately identified with PET. Similarly, the current evidence indicates that  $^{18}\text{F}$ -FDG PET is not useful for staging locoregional lymph nodes—especially the axillary nodes.

*Status of the Evidence.* The panel identified 5 systematic reviews (38,63,89–91), which analyzed 1, 10, 11, 13, and 15 studies. These systematic reviews included studies up to 1999, 2000, and 2004. The overall quality of the evidence was moderate, the quality of the systematic reviews themselves was moderate, and the quality of primary research evidence was moderate. No randomized studies were performed in this setting. The reviews did not make clear how many studies were retrospective and how many were prospective and enrolling consecutive patients. Many studies might have suffered from verification, detection, and spectrum biases. It was also not clear whether the same studies were included in more than 1 review. However, the findings among the reviews and the studies themselves were consistent. All systematic reviews concluded that, in general, sensitivity and specificity in detecting metastases are higher for PET than for conventional imaging tests. For example,

Prichard et al. (92) showed that the overall sensitivity and specificity of PET were 91% and 94%, respectively, compared with 57% and 45%, respectively, for CT. However, PET had a lower sensitivity and specificity than did SNB in the detection of lymph node metastases and cannot replace SNB for that use. PET also had a lower sensitivity than did CT in the detection of lung metastases. One systematic review (90) reported a 22% overall change-in-management rate with PET.

The panel concluded that adding  $^{18}\text{F}$ -FDG PET to conventional imaging is beneficial in detecting metastases in melanoma patients. The greatest benefit was attributed to extranodal metastasis detection, which, in turn, should help physicians administer treatment appropriate to the level of disease.

*Areas of Uncertainty and Implications for Future Research.* The role of  $^{18}\text{F}$ -FDG PET in the detection of melanoma is reasonably well established. However, further high-quality studies in this setting will be useful. In particular, studies evaluating the impact of PET on decision-making and patient outcomes, such as survival, will be most useful.

## PANCREATIC CANCER

### Is $^{18}\text{F}$ -FDG PET Useful as an Addition to CT in Diagnosing Pancreatic Cancer?

*Background and Rationale.* Pancreatic cancer is the fourth most common cancer in men and women in the United States (93). Cancer of the exocrine pancreas has an overall survival rate of less than 4% (94). Curative therapy is restricted to patients with limited and resectable disease. Late onset of often-nonspecific symptoms explains why most patients present with advanced and nonresectable disease at primary diagnosis. Despite a battery of imaging tools and recent advances in CT and MRI, the differential diagnosis of pancreatic adenocarcinoma and chronic focal pancreatitis is still a challenge (95). Therefore, the introduction of  $^{18}\text{F}$ -FDG PET in the detection of pancreatic cancer could have an important impact on such patients, avoiding unnecessary biopsies and surgeries.

*Evidence Summary.* The panel concluded that  $^{18}\text{F}$ -FDG PET should be added to conventional imaging in selected patients whose conventional imaging findings are inconclusive. The panel found moderate evidence that adding PET to CT will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries.

*Status of the Evidence.* The panel identified 2 systematic reviews (93,96), which analyzed 17 studies. Nine studies were included in both systematic reviews. The overall quality of the evidence was moderate, the quality of the systematic reviews themselves was high, and the quality of primary research evidence was moderate. No randomized studies were performed in this setting. Most of the studies prospectively enrolled consecutive patients. The major deficiency was a failure to describe the PET test procedure

adequately. Orlando et al. (93) compared PET/CT with CT to distinguish benign from malignant lesions. The sensitivity and specificity of  $^{18}\text{F}$ -FDG PET were 71%–100% and 53%–100%, respectively. On the other hand, the sensitivity and specificity of CT alone were 53%–100% and 0%–100%, respectively. The ROC of PET when CT findings were positive was 0.94, the sensitivity was 92%, and the specificity was 68%. The ROC of PET when CT findings were negative was 0.93, the sensitivity was 73%, and the specificity was 86%. The ROC of CT alone was 0.82, the sensitivity was 81%, and the specificity was 66%.

The BCBSA review (96) compared the use of PET with the use of a conventional work-up—including CT, MRI, and ultrasonography—and  $^{201}\text{Tl}$  SPECT. In all 9 studies, PET was superior. The unweighted pooled sensitivity of PET was 91%, and the pooled specificity was 86%. The PPV was 92% and the NPV was 84%, based on a prevalence of 66% derived from the included studies. The calculated LR+ and LR– were 6.5 and 0.15, respectively. Disagreement rates between PET results and conventional imaging results were 13%–54%. This group (96) also addressed the general question of whether the use of PET would alter patient management. They found 5 studies; however, they could make no conclusions because the results of these studies differed considerably.

The panel concluded that PET benefits the differentiation of benign from malignant lesions in the diagnostic work-up of patients with suspected pancreatic lesions. The greatest benefit was attributed to the possibility of excluding cancer without the need for biopsy or surgery, which may increase morbidity.

*Areas of Uncertainty and Implications for Future Research.* The available evidence is considerably consistent. Therefore, further research will probably simply confirm the current findings. The widespread use of new technology such as integrated PET/CT will likely further improve the diagnostic accuracy of  $^{18}\text{F}$ -FDG PET in this setting.

## SARCOMAS

### Is $^{18}\text{F}$ -FDG PET Useful for Diagnosing and Staging Sarcomas?

*Background and Rationale.* Sarcomas compose only 1% of all malignancies, and diagnosis and management are still a challenge (95). Soft-tissue sarcomas present with varied radiologic appearances.  $^{18}\text{F}$ -FDG PET has recently made promising contributions to patient management by providing important biologic information about soft-tissue malignant tumors and a noninvasive way to evaluate tumor metabolism (95). The use of PET in diagnosing and staging sarcomas might be helpful.

*Evidence Summary.* The panel concluded that the evidence is insufficient to support the use of  $^{18}\text{F}$ -FDG PET for diagnosing and staging sarcomas.

*Status of the Evidence.* The panel identified 1 systematic review (97), which analyzed 29 studies enrolling 5–202 patients per study. The overall quality of the evidence was

low. The authors of the systematic review used a checklist to assess the quality of the primary research studies and found it to be low. The quality of the systematic review itself was high, and the quality of primary research evidence was low. Ten studies evaluated only the detection of sarcomas, 10 evaluated detection combined with grading, 4 evaluated only grading, 5 evaluated therapy response, and 7 compared PET with another reference test (usually histopathology). The metaanalysis of pooled data on the detection of sarcomas (17 studies,  $n = 1,163$ ) yielded a sensitivity of 91% (95% CI, 89%–93%), a specificity of 85% (95% CI, 82%–88%), and a diagnostic accuracy of 88% (95% CI, 86%–90%). The calculated LR+ and LR– were 6.07 and 0.11, respectively. Ten studies contained data about mean standardized uptake value and concluded that the difference in that value between sarcomas and benign tumors was statistically significant; however, no cutoff value was described. In addition, the difference in mean standardized uptake value between low-grade and high-grade sarcomas was statistically significant for all studies and for mixed sarcomas. However, this difference was not statistically significant when the authors analyzed only soft-tissue sarcoma (again, no cutoff value was described). The authors stated that the data in these studies were insufficient to evaluate the role of PET in treatment response.

The panel concluded that misclassification of PET findings can potentially understage or overstage disease, resulting in less-than-optimal treatment for the actual level of disease or an increase in treatment-related toxicity.

*Areas of Uncertainty and Implications for Future Research.* Further research is likely to alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use. High-quality studies are needed in this setting, particularly studies evaluating the effect of a change in management on patient outcomes (e.g., survival or disease-free survival) in sarcoma.

## THYROID CANCER

### Is $^{18}\text{F}$ -FDG PET Useful for Detecting Recurrence of Thyroid Cancer?

*Background and Rationale.* Carcinoma of the thyroid gland is uncommon but is the most common malignancy of the endocrine system (98). Differentiated tumors (papillary or follicular) are highly treatable and usually curable. However, in approximately 10%–30% of patients thought to be disease-free after initial treatment, recurrence or metastases will develop. Because, on average, 25% of the recurrences cannot be detected by  $^{131}\text{I}$  whole-body scintigraphy, PET might have a potential role in detecting the recurrence of thyroid cancer.

*Evidence Summary.* The panel concluded that  $^{18}\text{F}$ -FDG PET should routinely be performed on patients previously treated for well-differentiated (follicular or papillary) thyroid cancer when the findings of  $^{131}\text{I}$  whole-body scintigraphy are negative and the thyroglobulin serum marker is more than 10 ng/mL. The probability of positive  $^{18}\text{F}$ -FDG PET findings increases with increasing levels of thyroglobulin and thyroid-

stimulating hormone. The panel found low-quality but consistent evidence that PET will likely improve important health-care outcomes in this setting and concluded that  $^{18}\text{F}$ -FDG PET is beneficial. This benefit is mostly due to avoiding futile surgeries and to identifying surgically resectable disease (as opposed to blind treatment of an elevated thyroglobulin level). However, the panel concluded against the use of  $^{18}\text{F}$ -FDG PET in the surveillance of thyroid cancer patients. No evidence supports the use of  $^{18}\text{F}$ -FDG PET when both the findings of  $^{131}\text{I}$  whole-body scintigraphy and the thyroglobulin serum marker are negative.

*Status of the Evidence.* The panel identified 2 systematic reviews addressing detection of recurrence. Hoof et al. (99) addressed the following issues:

- The diagnostic accuracy of  $^{18}\text{F}$ -FDG PET in detecting recurrence of follicular and papillary thyroid cancer: The overall quality of the evidence was low, the quality of the systematic review itself was high, and the quality of primary research evidence was low. This analysis included 14 studies: 7 prospective, 5 retrospective, and 2 of unclear design. Only 6 studies referred to inclusion of consecutive patients. The major deficiencies were selection, spectrum, verification, attrition, and detection biases. Also, the doses of  $^{131}\text{I}$  used for whole-body scintigraphy varied, and thyroid-stimulating hormone levels varied between the studies—both factors that may affect sensitivity and specificity. These studies analyzed 10–58 patients per study, and the number of patients totaled 402. Sensitivity was 70%–95% (data from 7 studies), and specificity was 77%–100% (data from 6 studies). In these 6 studies, PPV was 78%–100% and NPV was 68%–91%, with the prevalence of disease being 39%–90%.
- The use of  $^{18}\text{F}$ -FDG PET when the findings of  $^{131}\text{I}$  whole-body scintigraphy are negative and serum markers are elevated: To address this question, data from 156 patients (11 studies) were used. The overall quality of the evidence was low, the quality of the systematic review itself was high, and the quality of primary research evidence was low. The major deficiencies were selection, spectrum, verification, attrition, and detection biases. Verification of patient-level data was adequate in only 52% (68/131) of patients, 90% of whom proved to have recurrent disease. On average, in 82% (115/140) of the patients with raised markers and negative findings on  $^{131}\text{I}$  whole-body scintigraphy, PET localized foci of increased  $^{18}\text{F}$ -FDG uptake that were suggestive of recurrent disease.
- The use of  $^{18}\text{F}$ -FDG PET when the findings of  $^{131}\text{I}$  whole-body scintigraphy are negative and serum markers are not elevated: To address this question, the authors analyzed data from 5 studies, which enrolled a total of 50 patients. The overall quality of the evidence was low, the quality of the systematic review itself was high, and the quality of primary research evidence

was low. The major deficiencies were verification and detection biases. PET findings were negative in 34 (68%) of 50 patients. Using follow-up (1 y) and histology as reference tests, there was a false-negative result in 1 patient. However, false-positive findings, as documented by histopathology, seemed to be more frequent in this group than in the group of patients with raised serum markers.

- The use of  $^{18}\text{F}$ -FDG PET in comparison with other imaging modalities: The authors found 3 studies that compared  $^{18}\text{F}$ -FDG PET with  $^{99\text{m}}\text{Tc}$ -sestamibi,  $^{99\text{m}}\text{Tc}$ -tetrofosmin, or  $^{99\text{m}}\text{Tc}$ -furifosmin scintigraphy. These studies enrolled 20–54 patients. The overall quality of the evidence was moderate, the quality of the systematic review itself was high, and the quality of primary research evidence was moderate. Two of the 3 studies had a valid design (i.e., prospective, with masked interpretation and with the studies performed independently of other test results on each patient). Regarding individual tumor sites, both the  $^{18}\text{F}$ -FDG PET and the  $^{99\text{m}}\text{Tc}$ -sestamibi results were positive in 65% of patients,  $^{18}\text{F}$ -FDG was positive and  $^{99\text{m}}\text{Tc}$ -sestamibi negative in 25%, and  $^{18}\text{F}$ -FDG negative and  $^{99\text{m}}\text{Tc}$ -sestamibi positive in 10%.  $^{18}\text{F}$ -FDG PET images were of better quality and showed more lesions than did the  $^{99\text{m}}\text{Tc}$ -tetrofosmin images (135  $^{18}\text{F}$ -FDG-positive vs. 61  $^{99\text{m}}\text{Tc}$ -tetrofosmin-positive lesions). However, verification of  $^{18}\text{F}$ -FDG-positive/ $^{99\text{m}}\text{Tc}$ -tetrofosmin-positive and  $^{131}\text{I}$ -negative lesions was not performed. Overall, using patient-level data, PET had a sensitivity of 72% and a specificity of 100%, whereas  $^{99\text{m}}\text{Tc}$ -furifosmin imaging had a sensitivity of 33% and a specificity of 100%.
- The use of  $^{18}\text{F}$ -FDG PET in patients with known neoplastic foci: The authors reported that only 6 of the 14 reviewed studies contained data on  $^{18}\text{F}$ -FDG PET in patients with otherwise established recurrent disease. According to the inclusion criteria, only 1 study specifically addressed this patient group and no outcomes were described. Therefore, the authors concluded that the data were insufficient to draw any conclusion.

The other systematic review that addressed recurrence of thyroid cancer was performed by Facey et al. (38), who addressed the question of detecting recurrent disease in previously treated patients suspected of having metastatic disease on the basis of elevated serum markers and negative  $^{131}\text{I}$  whole-body scintigraphy findings. This systematic review included 11 studies, which enrolled a total of 244 patients. Of the 11 studies, 6 were included by Hoof et al. (99) in their systematic review. The overall quality of the evidence was low, the quality of the systematic review itself was low, and the quality of primary research evidence was unclear. Major deficiencies included substantial heterogeneity between studies, inconsistency in the definitions of

recurrence and cure between studies, and poor reporting of statistics for recurrence and cure. Overall, 65% of the cases were papillary cancer and 35% were follicular cancer. PET sensitivity was 84% (95% CI, 73%–91%), and specificity was 56% (95% CI, 27%–82%). The calculated LR+ and LR– were 1.91 and 0.29, respectively. Seven studies mentioned some change-in-management data, but further details were not provided.

Facey et al. also attempted to address the detection of recurrent medullary thyroid cancer in previously treated patients who had metastatic disease suspected on the basis of elevated serum markers and negative imaging findings. However, the data were insufficient to draw any conclusion. Only 6 studies, including 17 patients, were found.

The panel concluded that using  $^{18}\text{F}$ -FDG PET is beneficial in patients previously treated for thyroid cancer when the findings of  $^{131}\text{I}$  whole-body scintigraphy are negative and the level of thyroglobulin serum marker is elevated. The greatest benefit was attributed to detection of recurrence, which generally should help physicians administer treatment appropriate to the level of disease. Benefits are derived both from avoiding futile surgeries and from identifying surgically resectable disease as opposed to blind treatment based solely on elevated thyroglobulin levels. However, the use of  $^{18}\text{F}$ -FDG PET is not beneficial when both the  $^{131}\text{I}$  whole-body findings and the thyroglobulin serum markers are negative.

*Areas of Uncertainty and Implications for Future Research.* Further research is likely to alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use and will likely change the estimate. High-quality studies are needed in this setting.

## UNKNOWN PRIMARY TUMOR

### Is $^{18}\text{F}$ -FDG PET Useful for Detecting Unknown Primary Tumors?

*Background and Rationale.* An unknown primary tumor is defined as a biopsy-proven malignancy whose anatomic origin remains unidentified after diagnostic evaluation (100). The estimated incidence of unknown primary tumors is 2%–7% of all malignancies (100). Normally, unknown primary tumors are characterized by a poor prognosis, with a typical survival rate of no longer than 1 y from the time of diagnosis (101). Only 20%–27% of primary tumors are identified on conventional radiologic imaging (101). With the introduction of PET, the number of unknown primary tumors identified might increase, potentially improving patient outcomes by enabling administration of cancer-specific treatments.

*Evidence Summary.* The panel concluded that  $^{18}\text{F}$ -FDG PET should routinely be added to the conventional work-up of patients with unknown primary cancer. The panel found low but consistent evidence that the addition of PET to conventional imaging will likely improve important health-care outcomes and concluded that  $^{18}\text{F}$ -FDG PET is beneficial.

*Status of the Evidence.* The panel identified 2 systematic reviews (102,103), which analyzed 15 and 16 studies enrolling a total of 298 and 302 patients, respectively. The overall quality of the evidence was low, the quality of the systematic reviews themselves was high to moderate, and the quality of primary research evidence was low. No randomized studies were performed in this setting. The reviews did not make clear how many studies were retrospective and how many were prospective and enrolling consecutive patients. The major deficiencies were a failure to properly follow up patients and the small sample sizes of the studies. Delgado-Bolton et al. (102) showed that the sensitivity and specificity of PET for detecting primary tumor were 87% (95% CI, 81%–92%) and 71% (95% CI, 64%–78%), respectively. The reported LR+ was 3.05 (95% CI, 2.4–3.9), indicating that a positive PET result induced a small change in the pretest probability of detecting unknown primary tumors. The LR– was 0.17 (95% CI, 0.11–0.27), indicating that a negative PET result induced a moderate change in the pretest probability of detecting unknown primary tumors. Overall, PET detected 43% of the tumors (range, 35%–49%). Localization of the unknown primary tumor was described in 129 patients, and in 54 the lung was the site of the primary tumor.

The other systematic review (103) evaluated the use of  $^{18}\text{F}$ -FDG PET for detecting primary tumors in patients with cervical lymph node metastases after conventional imaging tests. In all studies, the conventional work-up was panendoscopy, CT or MRI, or chest radiography. The sensitivity of PET was 88%, the specificity was 75%, and the diagnostic accuracy was 79%.  $^{18}\text{F}$ -FDG PET detected 25% (74 patients) of tumors that were not apparent after the conventional work-up. In 24 patients, the primary tumor was detected both on PET and on MRI or CT. Finally,  $^{18}\text{F}$ -FDG PET led to the detection of previously unrecognized metastases in 27% of patients. The false-positive rate was 39% for the tonsillar area, 21% for the base of the tongue, and 8% for the hypopharynx.  $^{18}\text{F}$ -FDG PET had a lower sensitivity in detecting tumors at the base of the tongue (81%). Six studies (150 patients) provided change-in-management outcomes. PET was responsible for a therapeutic change in 25% of patients. In 111 patients, localization of the unknown primary tumor was described. Head and neck cancer was found in 84 patients, lung cancer in 20, and other types of cancer in the remaining 7.

The panel concluded that using PET in the diagnostic work-up of patients with unknown primary cancer is beneficial. The greatest benefit was attributed to detection of primary tumors that were not identified by the conventional work-up. This benefit, in turn, should help physicians administer appropriate, cancer-specific treatments to potentially improve survival.

*Areas of Uncertainty and Implications for Future Research.* Although the available evidence is consistent, further research is likely to alter confidence in the effect of  $^{18}\text{F}$ -FDG PET for this use. High-quality studies are needed



to assess the impact that cancer-specific treatments based on PET findings will have on the survival of patients with unknown primary tumors.

## SUMMARY

### Breast Cancer

*Diagnosis.* The panel concluded against routine use of  $^{18}\text{F}$ -FDG PET in the diagnosis of breast cancer but suggested use in specific clinical circumstances (e.g., high-risk patients with masses  $>2$  cm or aggressive malignancy and serum tumor marker elevation). The panel found moderate-quality evidence against routine use and concluded that the possibility of missing early-stage lesions and the high risk of false-negative results may be detrimental.

*Assessing Axillary Involvement.* The panel concluded against routine use of  $^{18}\text{F}$ -FDG PET for axillary staging of breast cancer. The panel found moderate-quality evidence that the use of PET for this purpose will likely misclassify the extent of breast cancer.

*Detecting Metastatic or Recurrent Breast Cancer.* The panel concluded that  $^{18}\text{F}$ -FDG PET should routinely be added to the conventional work-up for detecting metastatic or recurrent breast cancer in patients clinically suspected of metastasis or recurrence. The panel found moderate-quality evidence that  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries.

### Colorectal Carcinoma

*Diagnosis.* The panel concluded against routine use of  $^{18}\text{F}$ -FDG PET for detecting primary colorectal carcinoma. The panel found little evidence to support the use of  $^{18}\text{F}$ -FDG PET for this indication.

*Evaluating Hepatic Metastases.* The panel concluded that  $^{18}\text{F}$ -FDG PET should routinely be added to conventional imaging in preoperative diagnostic evaluations of patients with potentially resectable hepatic metastases of colorectal cancer. The panel found moderate-quality evidence that the use of PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries.

*Detecting Extrahepatic Recurrence or Local Relapse.* The panel concluded that PET should routinely be performed after the conventional work-up, especially if carcinoembryonic antigen levels are increased and the conventional work-up findings are negative. The panel found moderate-quality evidence that the use of PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries.

### Esophageal Cancer

The panel concluded that PET should routinely be performed as an additional tool for staging esophageal cancer.  $^{18}\text{F}$ -FDG PET can be particularly useful as an additional tool for detecting distant metastases, but its accuracy for detecting

local nodal metastases is still modest. On balance, indirect evidence supports the use of  $^{18}\text{F}$ -FDG PET in preoperative staging of esophageal cancer. The panel found moderate-quality evidence that the use of  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries.

### Head and Neck Cancer

The panel concluded that PET should be added to the imaging tests routinely used to identify unknown-primary head and neck tumors. The panel found moderate-quality but consistent evidence that adding  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial in this context.

*Diagnosis.* The panel concluded against routinely adding PET to CT or MRI in the diagnostic work-up of primary-tumor head and neck malignancies. The quality of evidence was insufficient to allow confident judgment on whether PET can determine the anatomic extent of primary head and neck malignancies to the level of certainty required for surgical resection.

*Staging.* The panel concluded that PET should routinely be added to CT or MRI to improve nodal or distant disease staging of head and neck cancer for the particular clinical circumstance. The panel found moderate-quality evidence that adding PET to CT or MRI will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial.

*Detecting Recurrence.* The panel concluded that PET should routinely be added to conventional imaging in the diagnostic work-up of potential recurrences of head and neck cancer. The panel found moderate-quality evidence that PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial.

### NSCLC

*Differentiating Benign from Malignant Lesions (Evaluation of SPN).* The panel concluded that PET should routinely be performed in the diagnostic work-up of patients with SPN. The panel found moderate-quality evidence that the recommended intervention will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding unnecessary surgeries in low-risk patients and enabling curative surgeries in high-risk patients.

*Staging.* The panel concluded that PET should routinely be added to the conventional work-up of NSCLC patients. The panel found high-quality evidence that PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries.

*Detecting Distant Metastases.* The panel concluded that PET should be obtained in the diagnostic work-up of lung cancer patients for distant metastases. The panel found moderate-quality evidence that  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that

the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries.

### **SCLC**

The panel concluded that the evidence is insufficient, of poor quality, or too inconsistent to conclude for or against PET in the management of SCLC.

### **Lymphoma**

*Staging.* The panel suggests that PET should routinely be added to the conventional work-up in pretreatment staging of lymphoma. The panel found low-quality but consistent evidence that adding  $^{18}\text{F}$ -FDG PET will improve important health-care outcomes and judged that the use of  $^{18}\text{F}$ -FDG PET is beneficial in this setting.  $^{18}\text{F}$ -FDG PET is considered more valuable in HD and early-stage aggressive NHL but less valuable in indolent NHL.

*Evaluating Bone Marrow Infiltration.* The panel concluded that PET may be added to bone marrow biopsy for staging and restaging lymphoma. The panel found moderate-quality evidence that adding  $^{18}\text{F}$ -FDG PET will improve important health-care outcomes and judged that the use of  $^{18}\text{F}$ -FDG PET is probably beneficial in this setting.

*Restaging or Detecting Relapse and Assessing Residual Mass or Progression After Completion of Initial Treatment.* Regarding HD, the panel concluded that in addition to the conventional work-up for restaging or detecting recurrence, PET should routinely be performed on patients to whom curative treatment was administered. The panel found moderate-quality evidence that adding  $^{18}\text{F}$ -FDG PET in this clinical setting will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial.

Regarding NHL, the panel concluded that in addition to conventional imaging for restaging or detecting recurrence, PET should routinely be performed on patients to whom treatment was applied with curative intent. The panel found low-quality but consistent evidence that adding  $^{18}\text{F}$ -FDG PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial.

Regarding follow-up of asymptomatic HD or NHL, the panel concluded against routine administration of PET for detecting relapse. The panel found limited evidence supporting the use of  $^{18}\text{F}$ -FDG PET in the routine follow-up of asymptomatic patients. A negative result would not affect the follow-up strategy, and a positive result could easily be detrimental because of the high frequency of false-positive results.

### **Melanoma**

Regarding staging and detecting recurrent metastatic melanoma, the panel concluded that PET should routinely be added to conventional imaging. The panel found moderate-quality evidence that adding PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by helping tailor treatment toward the stage of disease.

### **Pancreatic Cancer**

Regarding diagnosis of pancreatic cancer, the panel concluded that PET should be added to conventional imaging for selected patients in whom conventional imaging findings are inconclusive. The panel found moderate-quality evidence that adding PET to conventional imaging will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries.

### **Sarcomas**

Regarding diagnosis and staging of sarcoma, the panel concluded that the evidence is insufficient to conclude for or against intervention.

### **Thyroid Cancer**

Regarding detection of recurrence of thyroid cancer, the panel concluded that PET should routinely be performed on patients previously treated for well-differentiated (follicular or papillary) thyroid cancer when the findings of  $^{131}\text{I}$  whole-body scintigraphy are negative and the thyroglobulin serum marker is elevated ( $>10$  ng/mL). The panel found low-quality but consistent evidence that PET will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial, mostly by avoiding futile surgeries. However, the panel concluded against the use of  $^{18}\text{F}$ -FDG PET in the surveillance of thyroid cancer; therefore, the use of  $^{18}\text{F}$ -FDG PET is not recommended when both  $^{131}\text{I}$  whole-body scintigraphy and the thyroglobulin serum marker are negative.

### **Unknown Primary Tumor**

The panel concluded that PET should routinely be added to the conventional work-up of patients with unknown primary tumor. The panel found low-quality but consistent evidence that adding PET to conventional imaging will likely improve important health-care outcomes and concluded that the use of  $^{18}\text{F}$ -FDG PET is beneficial.

## **CONCLUSION**

### **Implications for Practice**

$^{18}\text{F}$ -FDG PET has become an established modality in the management of many cancers. Nevertheless, many uncertainties about its use remain and need to be addressed. PET should never be used alone but should be considered a supplement to other imaging modalities. Because hybrid PET/CT systems have virtually replaced standalone PET systems in the United States, correlative CT images are concurrently acquired in conjunction with PET images. Positive results should generally be confirmed by biopsy.

### **Implications for Research**

Given the poor quality of many of the diagnostic trials, the panel has provided explicit advice about the optimal design and conduct of future PET studies. In particular, studies that will demonstrate not only the superior diagnostic

accuracy of PET but also its clinical utility should be of high priority.

## APPENDIX A

### Glossary

**Attrition bias:** a systematic difference between comparison groups in the loss of participants from the study. If substantial differences are seen in dropouts or withdrawals from the study or between study arms (>10%), the results from the study should be viewed with extreme caution.

**Detection bias:** a systematic difference in outcome assessment. A method to prevent detection bias is to mask those who will assess the outcome (to the results of the index and reference test). In other words, researchers should mask the PET readers to the findings of other imaging tests or to histopathology findings.

**Selection bias:** a systematic difference in characteristics between those who are selected for study and those who are not. Selection bias can occur if a test is ordered on the basis of certain patient characteristics and can be avoided through the use of a prospective consecutive patient series, randomization, and adequate allocation concealment.

**Spectrum bias:** bias that occurs when diagnostic test performance varies across patient subgroups and a study of the performance of that test does not adequately represent all subgroups. Spectrum bias can occur when the study population includes a varied clinical spectrum—for example, cancer patients with early- and late-stage disease—in the same group. This type of bias is typically seen when patients in an ambulatory setting are mixed with those managed in a tertiary academic institution.

**Verification bias (or work-up bias):** bias that occurs when disease status (e.g., the presence or absence of histopathologically confirmed cancer) is not determined in all subjects who are tested and when the probability of verification depends on the test result, other clinical variables, or both. When verification of disease status is more likely among patients with positive PET findings, a bias is introduced that can markedly increase the apparent sensitivity of the test and reduce its apparent specificity. This bias can be avoided if all patients with the index test (e.g., PET scan) undergo the gold standard test regardless of the results of the index test (i.e., whether it was positive or negative).

**Likelihood ratio:** the likelihood that a given test result would be expected in a patient with the target disorder, compared with the likelihood that that same result would be expected in a patient without the target disorder. The likelihood ratio for a positive result (LR+) tells how much the odds of the disease increase when a test result is positive. The likelihood ratio for a negative result (LR-) tells how much the odds of the disease decrease when a test result is negative.

**Sensitivity:** a measure of the ability of a test to correctly detect people with the disease. Sensitivity is the proportion

of people with the target disease who are correctly identified by the test and is calculated as the number with disease who have a positive test result divided by the number with disease.

**Specificity:** a measure of the ability of a test to correctly identify people who do not have the disease. Specificity is the proportion of people without the target disease who are correctly identified by the test. Specificity is the complement of the false-positive rate (which is  $1 - \text{specificity}$ ) and is calculated as the number without disease who have a negative test result divided by the number without disease.

**Positive predictive value:** the probability that a patient has the disease if the test result is positive. Positive predictive value is calculated as the number with a positive result who have disease divided by the number with a positive result.

**Negative predictive value:** the probability that a patient with a negative test result does not have the disease. Negative predictive value is calculated as the number with a negative result who do not have disease divided by the number with a negative result.

**Target disorder:** the disease in which experimental outcomes (i.e., the performance of a diagnostic test) are measured.

**Index test:** the diagnostic test whose performance is being measured (e.g., PET).

**Reference standard test (gold standard test):** the test whose result is used to determine the true state of the subject (e.g., histopathology used to determine a patient's true disease state).

## REFERENCES

1. Ido T, Wan CN, Casella V, et al. Labeled 2-deoxy-D-glucose analogs:  $^{18}\text{F}$  labeled 2-deoxy-2-fluoro-D-glucose, 2-deoxy-2-fluoro-D-mannose and  $^{14}\text{C}$ -2-deoxy-2-fluoro-D-glucose. *J Labelled Comp Radiopharm.* 1978;14:175-183.
2. Warburg O, Posener K, Negelein E. Über den stoffwechsel der carcinomzelle. *Biochem Z.* 1924;152:309-335.
3. Sweet WH. The use of nuclear disintegration in diagnosis and treatment of brain tumors. *N Engl J Med.* 1951;245:875-878.
4. Hoffmann EJ, Phelps ME, Mullani NA, Higgins CS, Ter-Pogossian MM. Design and performance characteristics of a whole-body transaxial tomograph. *J Nucl Med.* 1976;17:493-503.
5. Phelps ME, Hoffman EJ, Mullani NA, et al. Design considerations for a positron emission transaxial tomograph (PET III). *IEEE Trans Biomed Eng.* 1976;NS-23:516-522.
6. Cho ZH, Farukhi J. BGL as a potential scintillation detector in positron cameras. *J Nucl Med.* 1977;18:840-844.
7. Phelps ME, Huang SC, Hoffman EJ, Selin C, Sokoloff L, Kuhl DE. Tomographic measurement of local cerebral glucose metabolic rate in humans with [ $^{18}\text{F}$ ] 2-fluoro-2-deoxy-D-glucose: validation of method. *Ann Neurol.* 1979;6:371-388.
8. Fowler JS, Ido T. Initial and subsequent approach for the synthesis of  $^{18}\text{F}$ FDG. *Semin Nucl Med.* 2002;32:6-12.
9. Alavi A, Reivich M. Guest editorial: the conception of FDG-PET imaging. *Semin Nucl Med.* 2002;32:2-5.
10. Shreve PD, Anzai Y, Wahl RL. Pitfalls in oncologic diagnosis with FDG PET imaging: physiologic and benign variants. *Radiographics.* 1999;19:61-77.
11. Bakheet SM, Powe J. Fluorine-18-fluorodeoxyglucose uptake in rheumatoid arthritis-associated lung disease in a patient with thyroid cancer. *J Nucl Med.* 1998;39:234-236.
12. Bakheet SM, Powe J, Kandil A, Ezzat A, Rostom A, Amartej J. F-18 FDG uptake in breast infection and inflammation. *Clin Nucl Med.* 2000;25:100-103.

13. Bakheet SM, Saleem M, Powe J, Al-Amro A, Larsson SG, Mahassin Z. F-18 fluorodeoxyglucose chest uptake in lung inflammation and infection. *Clin Nucl Med*. 2000;25:273–278.
14. Jones HA, Donovan T, Goddard MJ, et al. Use of <sup>18</sup>F-DG-pet to discriminate between infection and rejection in lung transplant recipients. *Transplantation*. 2004;77:1462–1464.
15. Weisdorf DJ, Craddock PR, Jacob HS. Glycogenolysis versus glucose transport in human granulocytes: differential activation in phagocytosis and chemotaxis. *Blood*. 1982;60:888–893.
16. Field MJ, Lohr KN, eds. *Guidelines for Clinical Practice: from Development to Use*. Washington, DC: Institute of Medicine, National Academy Press; 1992:38.
17. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328:1490.
18. Djulbegovic B. Principles of research synthesis. In: Perry M, ed. *ASCO Educational Session, 39th Annual Meeting*. Washington, DC: ASCO; 2003:737–750.
19. Early Breast Cancer Trialists Collaborative Group. Introduction and methods section. *Treatment of Early Breast Cancer: Worldwide Evidence 1985–1990*. Oxford, U.K.: Oxford University Press; 1990. Section 3.5. Review of many trials, with no data-dependent omissions, to limit selection bias in assessment of treatment effects.
20. Knottneurus J. *The evidence base of clinical diagnosis: evaluation of diagnostic procedures*. *BMJ*. 2002;324:477–480.
21. Montori VM, Wilczynski NL, Morgan D, Haynes RB. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *BMJ*. 2005;330:68.
22. Mijnhout GS, Riphagen II, Hoekstra OS. Update of the FDG PET search strategy. *Nucl Med Commun*. 2004;25:1187–1189.
23. Haynes RB, Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. *BMJ*. 2004; 328:1040.
24. Hilden J. Evaluation of diagnostic tests: the schism. *Society for Medical Decision Making Newsletter*. 2004;16:5–6.
25. Pauker SG, Kassirer J. The threshold approach to clinical decision making. *N Engl J Med*. 1980;302:1109–1117.
26. Sox HC, Blatt MA, Higgins MC, Maraton MC. *Medical Decision Making*. Boston, MA: Butterworths; 1988:84–97.
27. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes BR. *Evidence-Based Medicine: How to Practice and Teach EBM*. New York, NY: Churchill Livingstone Inc; 1997:50–80.
28. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326:41–44.
29. Antoch G, Stattaus J, Nemat AT, et al. Non-small cell lung cancer: dual-modality PET/CT in preoperative staging. *Radiology*. 2003;229: 526–533.
30. Gould MK, Kuschner WG, Rydzak CE, et al. Test performance of positron emission tomography and computed tomography for mediastinal staging in patients with non-small-cell lung cancer: a meta-analysis. *Ann Intern Med*. 2003; 139:879–892.
31. Kim JH, Czernin J, Allen-Auerbach MS, et al. Comparison between <sup>18</sup>F-FDG PET, in-line PET/CT, and software fusion for restaging of recurrent colorectal cancer. *J Nucl Med*. 2005;46:587–595.
32. Lardiniois D, Weder W, Hany TF, et al. Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. *N Engl J Med*. 2003;348:2500–2507.
33. Schoder H, Yeung HW, Gonen M, Kraus D, Larson SM. Head and neck cancer: clinical usefulness and accuracy of PET/CT image fusion. *Radiology*. 2004; 231:65–72.
34. *Guide to Clinical Preventive Services: Report of the U.S. Preventive Services Task Force*. 3rd ed. Baltimore, MD: Williams and Wilkins; 2000.
35. Dennis MA, Parker SH, Klaus AJ, Stavros AT, Kaske TI, Clark SB. Breast biopsy avoidance: the value of normal mammograms and normal sonograms in the setting of a palpable lump. *Radiology*. 2001;219:186–191.
36. Scheidhauer K, Walter C, Seemann MD. FDG PET and other imaging modalities in the primary diagnosis of suspicious breast lesions. *Eur J Nucl Med Mol Imaging*. 2004;31(suppl 1):S70–S79.
37. *FDG Positron Emission Tomography for Evaluating Breast Cancer*. Chicago, IL: Blue Cross and Blue Shield Association; 2000.
38. Facey K, Bradbury I, Laking G, Payne E. Positron emission tomography (PET) imaging in cancer management. Ultra Rapid Review. Health Technology Assessment. National Health Service (NHS) Research & Development (R&D) Programme. Southampton, U.K.: NHS R&H Programme; July 2004.
39. *Effectiveness of Noninvasive Diagnostic Tests for Breast Abnormalities: Comparative Effectiveness Review No. 2 (Prepared by ECRI Evidence-Based Practice Center Under Contract No. 290-02-0019)* [computer program]. Rockville, MD: Agency for Healthcare Research; 2006.
40. Isasi CR, Moedel RM, Blafox MD, et al. A meta-analysis of FDG-PET for the evaluation of breast cancer recurrence and metastases. *Breast Cancer Res Treat*. 2005;90:105–112.
41. Jemal A, Tiwari RC, Murray T, et al. Cancer statistics, 2004. *CA Cancer J Clin*. 2004;54:8–29.
42. Fernandez FG, Drebin JA, Linehan DC, Dehdashti F, Siegel BA, Strasberg SM. Five-year survival after resection of hepatic metastases from colorectal cancer in patients screened by positron emission tomography with F-18 fluorodeoxyglucose (FDG-PET). *Ann Surg*. 2004;240:438–447.
43. *FDG Positron Emission Tomography in Colorectal Cancer*. Chicago, IL: Blue Cross and Blue Shield Association; 2000.
44. Kinkel K, Lu Y, Both M, Warren RS, Thoeni RF. Detection of hepatic metastases from cancers of the gastrointestinal tract by using noninvasive imaging methods (US, CT, MR imaging, PET): a meta-analysis. *Radiology*. 2002;224:748–756.
45. Bipat S, van Leeuwen MS, Comans EF, et al. Colorectal liver metastases: CT, MR imaging, and PET for diagnosis—meta-analysis. *Radiology*. 2005; 237:123–131.
46. Wiering B, Krabbe PF, Jager GJ, Oyen WJ, Ruers TJ. The impact of fluor-18-deoxyglucose-positron emission tomography in the management of colorectal liver metastases. *Cancer*. 2005;104:2658–2670.
47. Huebner RH, Park KC, Shepherd JE, et al. A meta-analysis of the literature for whole-body FDG PET detection of recurrent colorectal cancer. *J Nucl Med*. 2000;41:1177–1189.
48. Liberale G, Van Laethem JL, Gay F, et al. The role of PET scan in the preoperative management of oesophageal cancer. *Eur J Surg Oncol*. 2004; 30:942–947.
49. Mariette C, Finzi L, Fabre S, Balon JM, Van Seuningen I, Triboulet JP. Factors predictive of complete resection of operable esophageal cancer: a prospective study. *Ann Thorac Surg*. 2003;75:1720–1726.
50. von Rahden BH, Stein HJ. Staging and treatment of advanced esophageal cancer. *Curr Opin Gastroenterol*. 2005;21:472–477.
51. Kato H, Miyazaki T, Nakajima M, et al. The incremental effect of positron emission tomography on diagnostic accuracy in the initial staging of esophageal carcinoma. *Cancer*. 2005;103:148–156.
52. van Westreenen HL, Westertep M, Bossuyt PM, et al. Systematic review of the staging performance of <sup>18</sup>F-fluorodeoxyglucose positron emission tomography in esophageal cancer. *J Clin Oncol*. 2004;22:3805–3812.
53. Miller FR, Hussey D, Beeram M, Eng T, McGuff HS, Otto RA. Positron emission tomography in the management of unknown primary head and neck carcinoma. *Arch Otolaryngol Head Neck Surg*. 2005;131:626–629.
54. Jereczek-Fossa BA, Jassem J, Orecchia R. Cervical lymph node metastases of squamous cell carcinoma from an unknown primary. *Cancer Treat Rev*. 2004;30:153–164.
55. *FDG Positron Emission Tomography in Head and Neck Cancer*. Chicago, IL: Blue Cross and Blue Shield Association; 2002.
56. Nieder C, Gregoire V, Ang KK. Cervical lymph node metastases from occult squamous cell carcinoma: cut down a tree to get an apple? *Int J Radiat Oncol Biol Phys*. 2001;50:727–733.
57. Vermeersch H, Loose D, Ham H, Otte A, Van de Wiele C. Nuclear medicine imaging for the assessment of primary and recurrent head and neck carcinoma using routinely available tracers. *Eur J Nucl Med Mol Imaging*. 2003;30:1689–1700.
58. Bohuslavizki KH, Klutmann S, Sonnemann U, et al. *Laryngorhinootologie*. 1999;78:445–449.
59. Braams JW, Pruijm J, Kole AC, et al. Detection of unknown primary head and neck tumors by positron emission tomography. *Int J Oral Maxillofac Surg*. 1997;26:112–115.
60. Jungehulsing M, Scheidhauer K, Damm M, et al. 2[F]-fluoro-2-deoxy-D-glucose positron emission tomography is a sensitive tool for the detection of occult primary cancer (carcinoma of unknown primary syndrome) with head and neck lymph node manifestation. *Otolaryngol Head Neck Surg*. 2000;123: 294–301.
61. Safa AA, Tran LM, Rege S, et al. The role of positron emission tomography in occult primary head and neck cancers. *Cancer J Sci Am*. 1999;5:214–218.
62. Branstetter BF IV, Blodgett TM, Zimmer LA, et al. Head and neck malignancy: is PET/CT more accurate than PET or CT alone? *Radiology*. 2005;235: 580–586.
63. *Health Technology Assessment of Positron Emission Tomography (PET) in Oncology: A Systematic Review*. Institute for Clinical Evaluative Sciences.

- April 2004. Available at: [http://www.ices.on.ca/webpage.cfm?site\\_id=1&org\\_id=31&morg\\_id=0&gsec\\_id=0&item\\_id=2080](http://www.ices.on.ca/webpage.cfm?site_id=1&org_id=31&morg_id=0&gsec_id=0&item_id=2080). Accessed January 14, 2008.
64. Ost D, Fein AM, Feinsilver SH. Clinical practice: the solitary pulmonary nodule. *N Engl J Med*. 2003;348:2535–2542.
  65. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet*. 1999;354:99–105.
  66. Fischer BM, Mortensen J, Hojgaard L. Positron emission tomography in the diagnosis and staging of lung cancer: a systematic, quantitative review. *Lancet Oncol*. 2001;2:659–666.
  67. Gould MK, Maclean CC, Kuschner WG, Rydzak CE, Owens DK. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. *JAMA*. 2001;285:914–924.
  68. Herder GJ, van Tinteren H, Golding RP, et al. Clinical prediction model to characterize pulmonary nodules: validation and added value of <sup>18</sup>F-fluorodeoxyglucose positron emission tomography. *Chest*. 2005;128:2490–2496.
  69. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules: application to small radiologically indeterminate nodules. *Arch Intern Med*. 1997;157:849–855.
  70. Gould MK, Sanders GD, Barnett PG, et al. Cost-effectiveness of alternative management strategies for patients with solitary pulmonary nodules. *Ann Intern Med*. 2003;138:724–735.
  71. Detterbeck FC, Falen S, Rivera MP, Halle JS, Socinski MA. Seeking a home for a PET, part 1: Defining the appropriate place for positron emission tomography imaging in the diagnosis of pulmonary nodules or masses. *Chest*. 2004;125:2294–2299.
  72. Pieterman RM, van Putten JWG, Meuzelaar JJ, et al. Preoperative staging of non-small-cell lung cancer with positron-emission tomography. *N Engl J Med*. 2000;343:254–261.
  73. Birim O, Kappetein AP, Stijnen T, Bogers AJ. Meta-analysis of positron emission tomographic and computed tomographic imaging in detecting mediastinal lymph node metastases in nonsmall cell lung cancer. *Ann Thorac Surg*. 2005;79:375–382.
  74. Dwamena BA, Sonnad SS, Angobaldo JO, Wahl RL. Metastases from non-small cell lung cancer: mediastinal staging in the 1990s—meta-analytic comparison of PET and CT. *Radiology*. 1999;213:530–536.
  75. Toloza EM, Harpole L, McCrory DC. Noninvasive staging of non-small cell lung cancer: a review of the current evidence. *Chest*. 2003;123(1, suppl):137S–146S.
  76. van Tinteren H, Hoekstra OS, Smit EF, et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. *Lancet*. 2002;359:1388–1393.
  77. Viney RC, Boyer MJ, King MT, et al. Randomized controlled trial of the role of positron emission tomography in the management of stage I and II non-small-cell lung cancer. *J Clin Oncol*. 2004;22:2357–2362.
  78. Herder GJ, Kramer H, Hoekstra OS, et al. Traditional versus up-front [<sup>18</sup>F] fluorodeoxyglucose-positron emission tomography staging of non-small-cell lung cancer: a Dutch cooperative randomized study. *J Clin Oncol*. 2006;24:1800–1806.
  79. Herder GJ, Kramer H, Hoekstra OS, et al.; POORT Study Group. Traditional versus up-front [<sup>18</sup>F] fluorodeoxyglucose-positron emission tomography staging of non-small-cell lung cancer: a Dutch cooperative randomized study. *J Clin Oncol*. 2006;24:1800–1806.
  80. Friedberg JW, Chengazi V. PET scans in the staging of lymphoma: current status. *Oncologist*. 2003;8:438–447.
  81. Hutchings M, Eigtved AI, Specht L. FDG-PET in the clinical management of Hodgkin lymphoma. *Crit Rev Oncol Hematol*. 2004;52:19–32.
  82. Isasi CR, Lu P, Blaufox MD. A metaanalysis of <sup>18</sup>F-2-deoxy-2-fluoro-D-glucose positron emission tomography in the staging and restaging of patients with lymphoma. *Cancer*. 2005;104:1066–1074.
  83. *FDG Positron Emission Tomography in Lymphoma*. Chicago, IL: Blue Cross and Blue Shield Association; 2000.
  84. Pakos EE, Fotopoulos AD, Ioannidis JP. <sup>18</sup>F-FDG PET for evaluation of bone marrow infiltration in staging of lymphoma: a meta-analysis. *J Nucl Med*. 2005;46:958–963.
  85. Zinzani PL, Fanti S, Battista G, et al. Predictive role of positron emission tomography (PET) in the outcome of lymphoma patients. *Br J Cancer*. 2004;91:850–854.
  86. Burton C, Ell P, Linch D. The role of PET imaging in lymphoma. *Br J Haematol*. 2004;126:772–784.
  87. Tsao H, Atkins MB, Sober AJ. Management of cutaneous melanoma. *N Engl J Med*. 2004;351:998–1012.
  88. Friedman KP, Wahl RL. Clinical use of positron emission tomography in the management of cutaneous melanoma. *Semin Nucl Med*. 2004;34:242–253.
  89. Mijnhout GS, Hoekstra OS, van Tulder MW, Teule GJ, Deville WL. Systematic review of the diagnostic accuracy of <sup>18</sup>F-fluorodeoxyglucose positron emission tomography in melanoma patients. *Cancer*. 2001;91:1530–1542.
  90. Schwimmer J, Essner R, Patel A, et al. A review of the literature for whole-body FDG PET in the management of patients with melanoma. *Q J Nucl Med*. 2000;44:153–167.
  91. Prichard RS, Dijkstra B, McDermott EW, Hill AD, O'Higgins NJ. The role of molecular staging in malignant melanoma. *Eur J Surg Oncol*. 2003;29:306–314.
  92. Prichard RS, Hill AD, Skehan SJ, O'Higgins NJ. Positron emission tomography for staging and management of malignant melanoma. *Br J Surg*. 2002;89:389–396.
  93. Orlando LA, Kulasingam SL, Matchar DB. Meta-analysis: the detection of pancreatic malignancy with positron emission tomography. *Aliment Pharmacol Ther*. 2004;20:1063–1070.
  94. Greenlee RT, Murray T, Bolden S, Wingo PA. Cancer statistics, 2000. *CA Cancer J Clin*. 2000;50:7–33.
  95. Zimny M. Diagnostic imaging of pancreatic cancer: the role of PET. *Front Radiat Ther Oncol*. 2004;38:67–75.
  96. *FDG Positron Emission Tomography in Pancreas*. Chicago, IL: Blue Cross and Blue Shield Association; 2000.
  97. Bastiaannet E, Groen H, Jager PL, et al. The value of FDG-PET in the detection, grading and response to therapy of soft tissue and bone sarcomas: a systematic review and meta-analysis. *Cancer Treat Rev*. 2004;30:83–101.
  98. Hundahl SA, Fleming ID, Fremgen AM, Menck HR. A National Cancer Data Base report on 53,856 cases of thyroid carcinoma treated in the U.S., 1985–1995. *Cancer*. 1998;83:2638–2648.
  99. Hoofst L, Hoekstra OS, Deville W, et al. Diagnostic accuracy of <sup>18</sup>F-fluorodeoxyglucose positron emission tomography in the follow-up of papillary or follicular thyroid cancer. *J Clin Endocrinol Metab*. 2001;86:3779–3786.
  100. Abbruzzese JLLR, Raber MN. Carcinoma of unknown primary. In: Abeloff MD, ed. *Clinical Oncology*. 2nd ed. Edinburgh, U.K.: Churchill Livingstone; 2000:2314–2329.
  101. Nanni C, Rubello D, Castellucci P, et al. Role of <sup>18</sup>F-FDG PET-CT imaging for the detection of an unknown primary tumour: preliminary results in 21 patients. *Eur J Nucl Med Mol Imaging*. 2005;32:589–592.
  102. Delgado-Bolton RC, Fernandez-Perez C, Gonzalez-Mate A, Carreras JL. Meta-analysis of the performance of <sup>18</sup>F-FDG PET in primary tumor detection in unknown primary tumors. *J Nucl Med*. 2003;44:1301–1314.
  103. Rusthoven KE, Koshy M, Paulino AC. The role of fluorodeoxyglucose positron emission tomography in cervical lymph node metastases from an unknown primary tumor. *Cancer*. 2004;101:2641–2649.
  104. Oxman AD. Checklists for review articles. *BMJ*. 1994;309:648–651.