Decision Memo for Positron Emission Tomography (FDG) (CAG-00065N)

Decision Summary

Conditions and limitations for coverage:

- We do not believe that it is reasonable and necessary to cover specific clinical indications for which adequate scientific data demonstrate that PET does not provide medical benefit. When such evidence exists, use in these indications will be specifically excluded from coverage.
- For use in oncologic diagnosis: PET is covered in clinical situations in which the PET results may assist in avoiding an invasive diagnostic procedure, or in which the PET results may assist in determining the optimal location to perform an invasive diagnostic procedure. PET is not covered for other diagnostic uses, and is not covered for screening (testing of patients without specific symptoms).
- For staging and restaging: Coverage for PET is subject to 2 conditions: 1) the stage of the cancer remains in doubt after completion of a standard diagnostic workup, including conventional imaging, and 2) clinical management of the patient would differ depending on the stage of the cancer identified. Use of PET would also be considered reasonable and necessary if it could potentially replace one or more conventional imaging studies.
- We consider restaging to include both restaging in the setting of recurrence and restaging following completion of a therapeutic regimen or to assess whether a complete response has been achieved. Use of PET to monitor tumor response during the planned course of therapy (i.e. when no change in therapy is being contemplated) is not covered.

Prior to obtaining an FDG PET study, the physician ordering this imaging procedure will be required to document in the patient's chart the specific clinical question that will be answered by the imaging study. The ordering physician will thereby be certifying the medical necessity of the study according to the conditions described above. This documentation is necessary in order for HCFA to be able to reliably review the appropriateness of use of FDG PET under the expanded coverage described in this document. HCFA plans to conduct a review within the first year following the effective date of this new coverage, and will use the results of this review to determine whether there is any need for further review and to decide whether revisions to the coverage policy would be indicated.

Back to Top

Decision Memo

This decision memorandum does not constitute a national coverage determination (NCD). It states CMS's intent to issue an NCD. Prior to any new or modified policy taking effect, CMS must first issue a manual instruction giving specific directions to our claims-processing contractors. That manual issuance, which includes an effective date, is the NCD. If appropriate, the Agency must also change billing and claims processing systems and issue related instructions to allow for payment. The NCD will be published in the Medicare Coverage Issues Manual. Policy changes become effective as of the date listed in the transmittal that announces the Coverage Issues Manual revision.

To: File: FDG Positron Emission Tomography (PET) CAG-00065N

From:

Sean Tunis, MD, MSc
Director, Coverage and Analysis Group

Mary Stojak
Health Insurance Specialist, Coverage and Analysis Group

Samantha Richardson
Health Insurance Specialist, Coverage and Analysis Group

Mitchell Burken, MD
Medical Officer, Coverage and Analysis Group

Michael Londner, MD, MPH
Medical Officer, Coverage and Analysis Group

Madeline Ulrich, MD
Medical Officer, Coverage and Analysis Group

Shana Olshan
Technical Advisor, Coverage and Analysis Group

Re: National Medicare Coverage Decision on FDG PET

Date: December 15, 2000

In this memorandum we: 1) describe FDG PET scans; 2) review the history of Medicare's coverage policy on PET scans and give an explanation of the coverage guidelines; 3) present and analyze the relevant scientific data including the literature submitted by the requestor; and 4) delineate the changed national coverage policy and HCFA's reason for the coverage decision policy. A summary of the new coverage policy can be found in the last section of this document prior to the appendices.

Description and Background of FDG Positron Emission Tomography (PET)

PET is a noninvasive diagnostic imaging procedure that assesses the level of metabolic activity and perfusion in various organ systems of the human body. Images are obtained from positron-emitting radioactive tracer substances (radiopharmaceuticals) that are usually administered intravenously to the patient

Positron-emitting radioisotopes were first discovered in the 1930's. FDG PET has been evaluated for several decades in pre-clinical models, and is premised on basic research in biochemisty and biology that have established the basis of glucose metabolism in normal cell function, and it's alteration in diseases like cancer, ischemic heart disease and some neurological disorders. The first PET scanners were developed in the United States in the 1970's with the first scan of a human reported in 1978. Through the early 1980's, PET scans were used primarily in research and predominantly focused on the neurosciences because scanners were typically only large enough for head studies. Due largely to the emergence of two major commercial suppliers in the mid-1980's, PET scanners have become capable of whole body imaging and increased computer processing capability. Improvements in the technology have had a significant impact on the quality of PET's image reconstruction and display.

PET's Ability to Identify Pathophysiology

Most of the disease-specific indications addressed in this coverage determination are related to PET use for various types of malignancies. As a group many of these diseases, which frequently are life-threatening, involve uncontrolled reproduction and spread of abnormal malignant cells. In adults, normal cells in most tissues divide only infrequently to replace worn-out or dying cells and to repair injuries. Malignant cells, which are both structurally and functionally abnormal, compete with and destroy normal cells and may spread throughout the body. They may aggregate in solid masses referred to as tumors. The spread of malignancy to a new site is called metastasis.

Classification of cancer by its appearance under a microscope and the part of the body in which it began, is important because different types of cancer vary in growth rates; how they spread through out the body, and in their susceptibility to various anticancer therapies. An accurate diagnosis of where the cancer originated in the body and its type is necessary so that the physician can determine the appropriate clinical management of the patient.

As a molecular diagnostic imaging modality, PET can detect rates of biological activity, as contrasted other imaging modalities such as x-ray films, computed tomography (CT), and magnetic resonance imaging (MRI), which depict the anatomical location of both normal and abnormal structures in the body. Malignancies can cause abnormalities of blood flow or metabolism before anatomic changes are apparent. Thus, disease can be detected by PET when anatomic imaging studies are still normal, and may be informative in differentiating benign from malignant processes. PET evaluation of tissue metabolism can indicate the probable presence or absence of malignancy based on observed differences of biologic activity, whereas anatomic imaging depends on the size and radiographic characteristics of lesions to determine the likelihood of malignancy. In addition, whole body imaging with PET provides the means to examine all organ systems for both primary and metastatic disease in a single procedure.

Safety of PET and Approval by the Food and Drug Administration (FDA) of FDG for PET Scans

The safety of PET is usually discussed in terms of the safety of the positron emitting radiopharmaceuticals or tracers. Silberstein (1998) conduced a study of 22 PET centers to determine what adverse reactions to the pharmaceuticals were observed retrospectively from the date the centers opened until 1994, and prospectively from 1994 to 1997. No negative effects were observed.

In 1972, FDA first approved a new drug application (NDA) for sodium fluoride F^{18} injection as a bone imaging agent to define areas of altered osteogenic activity. Marketing of this product ceased in 1975. Another tracer, Rubidium chloride 82 injection was approved in 1989 for assessing regional myocardial perfusion in the diagnosis and localization of myocardial infarction. The last tracer approved prior to 2000, was for the use of FDG injection for identification of regions of abnormal glucose metabolism associated with foci of epileptic seizures, in 1989.

On March 12, 2000, the FDA published a notice in the Federal Register that expanded approval of FDG for new indications. FDA concluded in that notice that a 10-millicuries (mCi) dosage (for adults) of FDG is safe and effective for oncological and cardiac applications. For cancer, FDG was specifically approved for assessing abnormal glucose metabolism to assist in evaluating malignancy in patients with known or suspected abnormalities found by other testing modalities or in patients with an existing diagnosis of cancer. This approval was based on 2 well designed studies of the use of FDG PET for specific oncologic applications, and 10 additional supporting studies of lower methodologic quality.

For cardiac applications, FDG was specifically approved for imaging of patients with coronary artery disease and left ventricular dysfunction, and when used together with myocardial perfusion imaging for identification of left ventricular myocardium with residual glucose metabolism and possible reversible loss of systolic function.

Summary of the History of Medicare's Coverage of PET Scans and an Explanation of the Coverage Guidelines

Medicare has reviewed the scientific literature regarding PET scans over a number of years, and has established coverage for six usesone in 1995, two in 1998, and three in 1999. All but the first use FDG as the tracer.

PET Scans using Rubidium 82 (Rb 82) for the Imaging of Perfusion of the Heart and Management of Patients with Known or Suspected Coronary Artery Disease

For services performed on or after March 14, 1995, Medicare first covered PET Scans using Rubidium 82 (Rb 82) done at rest or with pharmacological stress for the imaging of perfusion of the heart and management of patients with known or suspected coronary artery disease when:

- Used in place of, but not in addition to, a single photon emission computed tomography (SPECT), or
- The PET scan, whether rest alone or rest with stress, is used following a SPECT that was found inconclusive. In these cases, the PET scan must have been considered necessary in order to determine what medical or surgical intervention is required to treat the patient. (For purposes of this requirement, an inconclusive test is a test(s) whose results are equivocal, technically uninterpretable, or discordant with a patient's other clinical data.)

The coverage policy did not allow for PET scans using Rubidium 82 for the screening of asymptomatic patients, regardless of the number and severity of risk.

Staging of Non-Small Cell Lung Carcinoma (NSCLC)

Starting in January 1998, FDG PET scans were covered when used for the initial staging of suspected metastatic NSCLC in thoracic (mediastinal) lymph nodes in patients who have a confirmed primary lung tumor, but for whom extent of disease has not yet been established. The primary purpose of such staging is to determine the progress and extent of the disease, and based on that information to plan future management for the patient.

- Evidence of primary tumorA surgical pathology report is necessary to document the presence of an NSCLC.
- Whole body PET scan results and results of concurrent computed tomography (CT) and follow-up lymph node biopsyPET scans must be properly coordinated with other diagnostic modalities. The following reports are required to verify testing:
  1. the results of concurrent thoracic CT, which is necessary for anatomic information
  2. the results of any lymph node biopsy performed to finalize whether the patient will be a surgical candidate.

A lymph node biopsy is not covered in the case of a negative CT and negative PET, where the patient is considered a surgical candidate, given the presumed absence of metastatic NSCLC unless medical review supports a determination of medical necessity of a biopsy. A lymph node biopsy is covered in all other cases, i.e., positive CT+ positive PE T; negative CT+ positive PET; positive CT+ negative PET.

Coverage of FDG PET Scans for Characterization of Solitary Pulmonary Nodules

Also beginning in 1998, FDG PET scans were covered when used for the characterization of suspected solitary pulmonary nodules (SPNs). The primary purpose of such characterization should be to determine the likelihood of malignancy, in order to plan future management and treatment of the patient subject to the following conditions:

- Evidence of primary tumor Evidence of the initial detection of a SPN, usually by computed tomography (CT), is required.
- When other concurrent imaging techniques are also used, the results must be included on the claim.
- In the case of serial evaluation of SPNs using both CT and regional FDG PET chest scanning such PET scans will not be covered if repeated within 90 days following a negative PET scan.

In 1999, coverages of FDG PET for evaluation of recurrent colorectal cancer in patients with rising levels of carcinoembryonic antigen (CEA), for staging of lymphoma (both Hodgkin's and non-Hodgkin's) when the PET scan substitutes for a Gallium scan, and for the detection of recurrent melanoma were added.

Determining the Location of Recurrent Colorectal Tumors when Indicated by Rising Levels of CEA

In 1999, FDG PET was covered when used for determining the location of recurrent colorectal tumors when such tumors were indicated by rising levels of CEA. The use of FDG PET was limited to locating such tumors for the purpose of making a decision as to whether surgical intervention is warranted. However, the use of FDG PET to stage colorectal carcinoma was not covered under this national coverage decision. The provisions of the coverage policy were designed to limit coverage of PET to those situations in which it is effective in determining the course of future patient treatment. Determining the medical effectiveness of a service based on its utility in determining the course of treatment, is generally applied by Medicare to diagnostic modalities that are used as a substitute, or are intended to replace, other diagnostic modalities. The following conditions were also required:

- Evidence of documented previous colorectal carcinoma.
- Use of results of concurrent computed tomography (CT) and/or other diagnostic modalities when they are necessary for additional anatomic information.
- Frequency limitation of once every 12 months, unless medical necessity documentation supports a separate re-elevation of CEA within this period.

Staging of Lymphoma when Used as an Alternative to a Gallium Scan

Also determined in 1999, FDG PET scans became covered when used for staging lymphoma as an alternative to a Gallium scan when the following conditions are met.

- Evidence of disease Before the FDG PET scan is performed, a pathologic diagnosis of lymphoma must have already been made.
- When other concurrent imaging techniques are also used, the results must be included on the claim.
- Assurance that the FDG PET scan is an alternative to a Gallium scan.
- Limitation on use PET scans are not allowed any sooner than 50 days following the last PET scan or Gallium scan.
- Whole body FDG PET scans are covered only once every 12 months unless medical necessity documentation supports the specific need for localization of possible recurrent tumor within this period.

Evaluation of Recurrence of Melanoma Prior to Surgery and as an Alternative to a Gallium Scan

The last medical condition that became covered in 1999, was for the evaluation of melanoma prior to surgery in situations under the following conditions:

- Evidence of disease The patient must have previously been diagnosed with melanoma.
- When other concurrent imaging techniques are also used, the results must be included on the claim.
- Assurance that the PET scan is an alternative to a Gallium scan.
- Limitation on use – PET scans are allowed no sooner than 50 days following the last PET scan or Gallium scan.
- Full body PET scans are covered only once every 12 months unless medical necessity documentation supports the specific need for localization of possible recurrent tumor within this period.

Current FDG PET Scan Coverage Request

On July 10, 2000, HCFA received a request for broad coverage of FDG PET scans from Drs. Michael Phelps and Sam Gambhir. A list of 22 diseases was included in the request which covered various oncological conditions, myocardial viability, and neurological conditions. We determined that the appropriate benefit category fell under §1861(s)(3) diagnostic services. Due to volume of the evidence submitted by the PET community, we requested assistance from the Agency for Health Research and Quality (AHRQ). AHRQ had an Evidence-based Practice center (EPC) perform a validation check of the entire FDG PET submission.

The New England Medical center of Tufts University provided this validation. The EPC performed a literature search of the Medline and Biosis Previews databases for each of the clinical conditions listed in the PET request. The search was done to identify the universe of scientific evidence on the PET conditions submitted in the context of comparing the submitted material against a master bibliographic profile. The EPC conducted a search for potentially relevant PET scientific articles that dated from 1990 –2000. They located over 500 articles that were potentially relevant to the usage of PET scanners. The NEMC was not required to further analyze the data because HCFA did not request a full technology assessment.

The EPC concluded that the PET request was not presented as a standard systematic literature review, but represented a large bibliographic compilation of the literature. The NEMC report raised some questions about relevant studies that might not have been included in the PET request, and identified several errors in the data that were abstracted from the studies to create the summary tables. HCFA concluded that it would be necessary to conduct independent systematic reviews of the FDG PET literature in order to produce appropriate coverage policy.

In order to assure a full and open public discussion of the scientific and clinical issues raised by FDG PET, we requested advice from the Medicare Coverage Advisory Committee (MCAC) on October 17th. The Executive Committee of the MCAC met on November 7, 2000 to consider guidelines for the evaluation of diagnostic tests in general and to consider selected issues (i.e. colorectal cancer management, differential diagnosis of dementia, and lung cancer diagnosis and staging) from the PET coverage request.

After an overview of PET presented by Dr. Phelps, the Executive Committee chairman, Dr. Harold Sox, presented the Working Framework for Evaluating Diagnostic Tests, found in Appendix B. The Guidelines were discussed by the panelists, but not subjected to a formal vote.

It was the sense of the panel that one should consider first whether the evidence is sufficient to establish that a test under consideration provides diagnostic information that is at least as effective as standard alternatives. The Committee then made suggestions of issues future panels might want to consider in assessing the impact on health outcomes of particular diagnostic tests. Following public comments, the Committee discussed application of the guidelines to some of the proposed new uses for PET. Although there was no formal vote, generally, the Committee suggested that PET had benefit in assessing recurrent colorectal cancer and that there was some evidence that might be generalized to the use of PET in other applications. It was noted that the performance of PET may differ depending on the specific cancer being evaluated and the physical location of the cancer and any possible metastases. It was further suggested that the MCAC Diagnostics Panel might look into the details of extending coverage for other oncologic indications based on the evidence related to colorectal cancer. Some Committee members also expressed concern that HCFA's policy might prevent coverage for PET use for certain cancers because their rarity precluded performance of necessary studies.

Quality of Studies Evaluating Diagnostic Technology

Over the past decade, the characteristics of high quality studies for evaluating diagnostic tests have been well-documented in a number of committee reports and peer-reviewed publications. Most of these documents come to similar conclusions about the study design characteristics that are helpful in reducing bias, and ensuring that the reported results are an accurate reflection of the performance of the test. These characteristics are similar to those included in the following chart.

Experimental Design Features That Enhance Scientific Rigor of Diagnostic Test Evaluations*

Design Feature	Comments
Defining the problem and hypotheses	• Helps to clarify the clinical problem • Inclusion and exclusion criteria are defined to reduce confounding variables
Adequate patient sample size for sufficient statistical power	• Depends on the expected magnitude of effect and whether all patients have both competing imaging tests
Patient referral sources that include a clearly defined broad spectrum of disease presentation and severity	• Reduces referral bias (spectrum bias)**
Clearly defined patient groups based on pre-test probability estimates	• Allows reader to judge generalizability of findings to his/her practice • Offsets referral bias • Consider adequate sample size for each subgroup analysis
All patients have comparison tests and similar follow-up	• Reduces work-up bias***
Randomized, independent, blinded reading of competing tests	• Avoids test review bias**** • Consider blinding test interpreters to clinical information, other tests, and final diagnosis • Should develop methods to reduce interobserver variation
Expert interdisciplinary gold standard panel and determination of true diagnosis	• Diagnosis determined both with and without test results allow measurement of the degree of diagnostic review bias (incorporation bias)***** in result
Outcomes analysis	• Data on operating test characteristics are gathered using a research protocol • Data on consequences of diagnostic and treatment choices on patient outcomes are obtained from the literature

* Adapted from Veterans Health Administration Report (1997)
** referral bias relates to the differences among patient populations in the spectrum of disease presentation and severity
*** work-up bias most commonly occurs when results from one test determines inclusion or exclusion from the study or from further work-up
**** test review bias occurs when the final diagnosis or results of the comparison test are used in planning or interpreting the test under study
***** diagnostic review bias occurs when the gold standard diagnosis is influenced by results of the imaging test

These principles of study design were incorporated into the Proposed Guidelines for Evaluating Diagnostic Tests discussed by the MCAC Executive Committee (see Appendix D). The following chart from that document illustrates how failure to follow established principles of scientific investigation can weaken study results.

Ideal study	Usual study	Effect of Usual Study
The study subjects are consecutive patients seen in a typical clinical setting with a chief complaint.	Subjects selected because they have had the diagnostic gold standard.	Overestimates sensitivity and underestimates specificity.
All patients who get the index test also get the reference test.	Patients with negative results on the index test often don't get the diagnostic gold standard.	Overestimates sensitivity and underestimates specificity.
The person who interprets the index test is blinded to all other information.	The person who interprets the index knows the clinical history and the results of the diagnostic gold standard.	Overestimates sensitivity and specificity.
The person who interprets the reference test is blinded to all other information.	The person who interprets the diagnostic gold standard knows the clinical history and the results of the index test.	Overestimates sensitivity and specificity.
The reference test is a valid measure of the disease state.	The diagnostic gold standard imperfectly measures the disease state.	The measured test performance could either be worse or better than the true performance.

Analysis of the Relevant Scientific Data

For this coverage request, we have supplemented the requestor's submission with technology assessments published in 2000 by Blue Cross Blue Shield, the Report of the Commonwealth Review of Positron Emission Tomography (also published in 2000), and additional analysis on lung and esophageal cancers using material from the requestor's submission. In the next section of this memorandum, we outline a number of disease-specific indications for use of FDG PET. Our coverage NCDs are based on the use of those assessments, the requestor's submission, and our limited literature review using the same basic questions that the MCAC used in their Working Diagnostic Guidelines.

In addition to the published data reviewed above, HCFA also considered other forms of evidence, including extensive consultation with clinical experts in oncology, nuclear medicine, cardiology, neurology, and other relevant clinical disciplines. We also took into account the basic biology and biochemistry of disease upon which PET imaging technology is based. All of this information was helpful in interpreting the direct empirical studies that have been performed to evaluate the test performance and clinical utility of PET. The relevant body of evaluation literature is briefly summarized in this section for the subset of clinical applications of PET addressed in this coverage decision memo.

Lung Cancer (Non-Small Cell)

Background

HCFA coverage has already been provided for evaluating solitary pulmonary nodules, as well as staging non-small cell carcinoma of the lung (NSCLC), making it logical to inquire about evidence which might support applying FDG PET to detecting residual or recurrent NSCLC. The submitted package by UCLA includes a relatively large diagnostic trial by Bury *et al.*(1999) which supports this application. In a group of 126 consecutive patients, divided into 58 who were in an early curative group and 68 in an early palliative group, there was considerably higher sensitivity for PET (100%) vs. CT (72%). Please note that specificities were both equivalent (>90%), and the same performance trends were found in each patient subgroup.

Application of working diagnostic guidelines:

- *Is this study of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

Other than a reported absence of blinding, there was no strong source of bias, given the relatively large sample size and use of consecutive patients to minimize selection bias.

- *What is the potential impact of PET accuracy upon health outcomes?*

Per the relatively strong study by Bury, one may surmise that PET provides an appropriate degree of diagnostic accuracy. However, subsequent outcomes data are not furnished.

Recommendation:	There is evidence to support the role of PET detecting residual or recurrent tumor after treatment of NSCLC.
Rationale:	The Bury study provides an evidentiary rationale for supporting this particular application of FDG PET.

Esophageal Cancer

Background

Esophageal cancer is a relatively rare but lethal type of cancer, which is newly diagnosed in approximately 10,000 Americans each year. This tumor has been considered synonymous with squamous cell carcinoma. However, adenocarcinoma is now more common in the United States, and has a rising incidence rate. Although the overall 5-year survival rate has remained steady at about 5%, patient management may be greatly assisted by diagnostic techniques that can properly assign patients to a curative subgroup, where extension of disease has not already disqualified patients for surgery. The 5-year survival rate with surgical intervention alone is approximately 30%.

Role of PET in Pre-Surgical Staging

The below scenarios can be described, whereby PET, if it is demonstrated to have added diagnostic benefit, could be used as an adjunct to conventional imaging (CI), such as computed tomography (CT) and ultrasound (US), in primary staging:

- If CI is negative and PET is positive for metastatic disease, then it is likely that the patient has unresectable disease and curative surgery is not applicable;
- If CI is negative and PET is negative, then, conversely, it is likely that the patient is a candidate for curative surgery;
- If CI is positive, then it may not be likely that PET is even needed since the patient has already been demonstrated to have unresectable disease.

In tandem with full-length articles provided by the PET request package, a supplemental Medline search (Ovid) was conducted for the textwords "esophageal cancer" and "PET," with limitations to human studies in English, published from 1997-2000. The following inclusion criteria were applied such that eight studies were selected for further review:

- Study sample included at least 10 patients;
- Patient sample homogeneous with respect to type of primary cancer, and
- Study described correlation of FDG PET findings with data from an appropriate reference standard, for at least some of the patients in its sample.

Three of these studies (Kole *et al.* 1998, Luketich *et al.* 1999, Flamen *et al.* 2000) report comparisons in the ability of CT to measure distant metastases versus FDG PET.

The Kole study combines factors for overall resectability, demonstrating an accuracy of 65% for CT versus 88% for PET ($p = 0.04$, using McNemar test), but without mention of sensitivity and specificity values. The Luketich study demonstrates a sensitivity of 69% and specificity of 93% for PET versus 46% and 74%, respectively for CT. Finally, Flamen corroborates this favorable trend, by reporting a PET sensitivity/specificity of 74%/90% in detecting Stage IV disease versus 47%/78% for CT plus US. Thus, in all three studies, there is evidence of PET's additional diagnostic benefit for assessing metastatic disease.

Four studies (Flanagan *et al.* 1997, Block *et al.* 1997, Luketich *et al.* 1997, Choi *et al.* 2000), in addition to Flamen and Kole, provided data on nodal evaluations, and there was at least comparable performance data for both conventional imaging and PET.

Application of working diagnostic guidelines:

- *Are the studies of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

Of the three key studies used to support the relative benefit of PET in detecting metastatic disease, the Flamen article reveals no significant sources of bias. Although both Kole and Luketich *et al.* 1999, both used consecutive patients, each study did not apply all diagnostic tests to all patients, and the latter study also failed to demonstrate blinding.

- *What is the potential impact of PET accuracy upon health outcomes?*

Yeung *et al.* 1999 and Flanagan have shown patient management changes of 14% and 17%, respectively, with respect to the use of PET. Furthermore, Luketich *et al.* 1999 used Kaplan-Meier survival analysis to demonstrate that patients with local disease on PET had a 30-month survival of 60% versus 20% survival for those who had distant disease on PET. These findings suggest that the use of PET has a positive effect upon health outcomes.

Recommendation:	Evidence is present to support the use of PET in pre-surgical staging of esophageal cancer.
Rationale:	Multiple studies provide the basis for such coverage. The clinical dilemma posed by the limitations of conventional imaging can be, in part, addressed by the further use of functional PET imaging.

Role of PET in Monitoring Recurrence

There is very limited data to suggest that PET can be a valuable tool for monitoring treatment; however, the Yeung study profiles 84/150 scans for this type of indication. Although there is combined data for both staging and recurrence in this data set, the overall superior performance of PET (80% sensitivity and 95% specificity) versus CT (68% and 81%, respectively) would suggest that some benefit may be conferred for this use of PET in managing esophageal cancer.

Recommendation:	Some evidence supports extension of PET coverage into this additional indication for esophageal cancer.
Rationale:	Unless strong negative evidence is present, HCFA can use this evidence to support broadening coverage within this tumor type.

Colorectal Cancer

Background

Carcinoma of the large bowel is by far the most common and most curable carcinoma of the gastrointestinal tract, with approximately 140,000 new cases per year and 55,000 deaths per year. Males and females are affected equally, the mean age of incidence is 62 years. Different stages of tumor have been classified which depend upon whether: The tumor involves the wall of the bowel only, there is extension through the wall, there is lymph node metastatic disease, or there is distant metastatic involvement. Therefore, multiple patient management checkpoints will be evaluated where FDG PET may contribute useful diagnostic information:

- The ability of PET to differentiate local recurrence of tumor from postoperative scarring at the primary surgical site;
- The role of PET to provide additional benefit over conventional imaging for primary staging of hepatic and extrahepatic disease, before any surgery/therapy has been undertaken, and
- The role of PET for assessing recurrent colorectal cancer beyond simply where the tumor marker carcinoembryonic antigen (CEA) serves as a trigger for investigation, noting that current HCFA policy allows for PET evaluation only in the context of a rising CEA.

Distinguishing Local Recurrence from Postoperative Scar

In patients who have undergone primary resection for colorectal cancer, FDG PET may be instrumental in detecting whether tumor has recurred at the surgical site. The following management alternatives are faced by patients who present with this dilemma:

- Biopsy the area, or

- Perform a test, such as a PET scan, which may reduce the probability that an indurated area is recurrent cancer, such that, in turn:
    - If the PET scan is negative, conduct watchful waiting, or
    - If the PET scan is positive, proceed to biopsy.

Beneficial outcomes (true negatives) occur when the PET scan correctly shows that a local lesion is a post-operative scar, and a biopsy procedure may be rendered unnecessary. Conversely, adverse outcomes (false negatives) occur when PET incorrectly suggests the area in question is postoperative scar, thus causing clinicians to forego biopsy which could have shown recurrent tumor. This incorrect imaging result could presumably result in a missed opportunity for curative resection.

Consequently, PET would demonstrate greater clinical utility based upon its ability to generate a very high negative predictive value (NPV) in which there is a relatively low proportion of false negative results. Therefore, in the context of a high NPV, a patient might elect to forego a tissue sampling procedure and continue with less invasive monitoring.

We chose six studies which were selected for review from the Blue Cross/Blue Shield TEC assessment, using the following inclusion criteria:

- Study published or accepted for publication as a full article in a peer-reviewed journal;
- Study sample included at least 10 patients;
- Patient sample homogeneous with respect to type of primary cancer;
- Study performed tomographic, not planar, imaging with FDG as the radiotracer, and
- Study described correlation of FDG PET findings with data from an appropriate reference standard, for at least some of the patients in its sample.

Even though there was a high sensitivity = 96% and high specificity = 98%, the Bayesian estimate of NPV was 92%, given the unweighted pooled probability of local recurrence = 69%. This pooled NPV estimate of 92% means that the probability of occult local recurrence in patients with negative PET scans is 8%. Please note that if this prevalence of local recurrence had only been 5%, given the same values of sensitivity and specificity, the NPV would have been much higher at 99.8%.

Application of working diagnostic guidelines:

- *Are the studies of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

The six studies provided useful diagnostic performance data, and this assertion was also confirmed by the MCAC Executive Committee panelists, subject to the following potential sources of bias:

- Consecutive patient enrollment was not required as a means of minimizing selection bias,
- none of the six studies successfully demonstrated blinding protocols,
- a gold standard reference test was not required for all study patients, and
- two out of six studies only had 15 and 18 patients, respectively.

- *What is the potential impact of PET accuracy upon health outcomes?*

The TEC assessment postulated that patients and their physicians would be unlikely to forego histologic sampling, based upon PET scan findings, with a false negative rate as high as 8%, since this could likely cancel/delay re-operation, which has an approximate 20% chance of cure. The MCAC panelists expressed similar concerns that this reported false negative rate would impose such a barrier. However, since the panelists also surmised that using PET for suspected local recurrence could, in turn, pick up additional extra-pelvic metastases (see Schiepers *et al.*1995), there was a majority opinion that PET imaging could have a favorable impact upon patient management/health outcomes. Thus, it may not be pertinent to consider PET scanning for recurrent tumor which occurs only at the local resection site, but PET would be helpful in detecting more widespread recurrent disease.

Recommendation:	Coverage is supported for FDG PET to help differentiate post-operative scar from the recurrence of colorectal carcinoma.
Rationale:	It appears that PET scanning has the ability to influence the post-test probabilities such that patients and their physicians can choose an appropriate biopsy strategy which, in turn, maximizes the opportunity for curative resection of recurrent colorectal carcinoma.

Detecting Hepatic and Extrahepatic Metastases

The detection of hepatic and extrahepatic metastases by clinicians can improve the selection of surgical candidates. Patients with non-resectable metastases can be more accurately identified, so that unnecessary surgery can be avoided.

The logic of the Blue Cross/Blue Shield TEC causal chain is as follows, assuming that PET follows conventional imaging (CI):

- If CI demonstrates resectable disease, with cure potentially achievable in approximately 30% of patients, then either:
    - CI and PET are concordant such that surgery is pursued, or
    - CI and PET are discordant such that palliation is chosen in lieu of surgery.

- If CI demonstrates non-resectable disease, then either:

- Concordance of CI and PET avoids unnecessary surgery, or
- Discordance of CI and PET encourages the pathway of curative surgery.

The previous July 1999 coverage instructions, which were issued after review of the presentations made at a January 1999 PET Town Hall Meeting, granted coverage for PET when recurrence is suspected as a result of a rising serum CEA level. Therefore, additional coverage deliberations on this issue should address the following two narrow questions:

1. Does PET provide additional benefit over CI in primary staging of hepatic and extrahepatic disease, before any surgery/therapy has been undertaken?
2. Should the assessment of recurrent colorectal cancer only be limited to situations where rising CEA serves as a trigger for investigation?

Only one study by Abdel-Nabi *et al.*(1998) presents data on primary staging. With respect to hepatic metastases, this study showed a sensitivity of 38% for CT, as opposed to 88% for PET, and specificities of 97% and 100% for CT and PET, respectively. Regarding extrahepatic nodal metastases, both CT and PET had a sensitivity of 29%, compared to specificities of 85% and 96% for CT and PET, respectively.

Application of working diagnostic guidelines for primary staging of metastatic lesions:

- *Is this study of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

Study strengths included consecutive recruitment of patients to minimize selection bias, Sample size nearly 50 (n = 48), and 44/48 patients subjected to a desirable surgical gold standard. The major study limitation was an unblinded design.

- *What is the potential impact of PET accuracy upon health outcomes?*

When presented to the MCAC panelists, the post-test probability data for both primary staging, as well as for assessment of recurrence, received a positive response with respect to patient management changes/improved health outcomes. Based upon this general acceptance of the studies, one may infer a positive response for the use of PET related to can primary staging.

Recommendation: Coverage is supported for the use of FDG PET when determining the presence of hepatic/extrahepatic metastases in the primary staging of colorectal carcinoma, prior to selecting a treatment regimen.

Rationale:

The relatively strong findings presented in the Abdel-Nabi study provide the evidentiary basis for this recommendation.

Application of working diagnostic guidelines for evaluating recurrent hepatic/extrahepatic disease when there are indicators other than rising CEA:

In 1999, Valk *et al.* presented data related to the issue of use of PET in the absence of rising CEA. Other studies have looked at rising CEA in combination with other indicators of suspected recurrence (e.g., abnormal CT scan). In the Valk study, a subgroup of 76 patients were referred for PET based solely upon positive CT findings (i.e., solitary recurrent lesion), as opposed to some admixture of rising CEA, CT, etc. PET results altered post-test probabilities in several ways:

- 47 patients had confirmed localized single recurrences with PET and proceeded to intended curative surgical follow-up;
- 23 patients were found to have unsuspected sites of recurrence, thus altering patient management, such that 10/23 patients did not undergo surgery; and
- 6 patients showed no tumor, causing 2 patients to defer surgery in favor of clinical follow-up (please note that both patients were free of recurrent tumor 14-32 months after PET scan).

- *Is this study of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

The Valk study's particular strengths is its recruitment of 155 consecutive patients, with relatively more complete blinding than many other PET studies.

- *What is the potential impact of PET accuracy upon health outcomes?*

The above data provide evidence that rising CEA should not be viewed as the only trigger for evaluating recurrent disease. Although the change in eventual health outcomes may not be obvious from this limited data, there were documented patient management changes as a result of PET imaging under this clinical scenario.

Recommendation:	Coverage is supported for expanding the role of evaluating recurrent hepatic/extrahepatic colorectal cancer beyond the limited presentation of a rising CEA level.
Rationale:	Whereas rising CEA provides the most obvious trigger for evaluating colorectal cancer recurrence, the ability to tease out other potential risk factors was limited by several studies in which rising CEA was combined with multiple other factors. However, the Valk study presents convincing data on abnormal CT scans which, in turn, support a less restrictive approach to monitoring recurrence of colorectal cancer.

Lymphoma

Staging and restaging of both Hodgkin's and non-Hodgkin's disease have previously been approved for Medicare coverage. The recent Blue Cross/Blue Shield TEC Assessment supported that determination.

Melanoma

Background

Malignant melanoma, which is a relatively aggressive cancer arising primarily in the skin, affected 44,000 new patients and resulted in 7,300 deaths in 1999 and this number continues to rise. Invasive melanoma is classified in four categories I – IV ranging from primary tumor, with thickness less than 1.5 mm, to extranodal metastastic involvement. These stages can be quantified to account for incidence of disease and mortality. Localized disease accounts for 82% of new disease and has a five-year survival of 87.7 % while distant disease accounts for 4% of new disease, but five-year survival is only 12.6%.

The review of PET with regard to melanoma will include two indications:

1. Detecting regional lymph node metastases in either initial staging or monitoring after primary treatment, and
2. Detecting extranodal metastasis at initial staging or during follow-up after treatment.

Detecting regional lymph node metastases during either initial staging or monitoring after primary treatment

It is essential to first emphasize that HCFA already covers monitoring after primary treatment; therefore, the current discussion is limited to the detection of regional lymph node metastases during initial staging.

This question addresses patients who have clinically localized disease with invasive cutaneous lesions of intermediate thickness (1.0-4.0 mm). For these patients, PET may be beneficial in determining the appropriateness of sentinel node biopsy (SNB). Traditionally, when patients are diagnosed with local disease, they undergo SNB to determine the need for elective lymph node dissection. If PET can be demonstrated to be as sensitive and specific as SNB in determining lymph node metastases, then these persons can be spared the possible adverse effects from SNB. When both PET and SNB are concordant, there is no change in management and no harm in a less invasive approach. The caveat arises when PET is falsely positive or negative. When PET is falsely negative, the patient forgoes or delays potentially beneficial lymph node dissection, and when PET is falsely positive, the patient undergoes an unnecessary dissection.

Thus, the required study design compares the accuracy of PET to SNB, and there are four possible outcomes:

- PET positive and SNB positive (concordant true positive): In this instance both studies would recommend elective lymph node dissection.
- PET negative and SNB negative (concordant true negative): In this instance both pathways direct the safe avoidance of lymph node dissection. In fact if PET were equal or better than SNB, PET would avoid SNB as well. This is precisely the group PET looks to impact.
- PET negative and SNB positive (discordant false negative): This is the dangerous category since patients with true disease would forgo or delay elective lymph node dissection.
- PET positive and SNB negative (discordant false positive): These patients would get over treated with elective lymph node dissection.

The following study selection criteria were used in the Blue Cross/Blue Shield TEC Assessment of this issue:

- Published or accepted for publication as a full article in a peer-reviewed journal;
- At least 10 patients;
- Patient sample homogeneous with respect to type of primary cancer (i.e., studies excluded if there were either patients with various tumor types or if there was a mixture of primary and metastatic lesions);
- Performed tomographic rather than planar imaging with FDG as the radiotracer, and
- Correlation of PET findings with data from an appropriate reference standard, for at least some of the patients in the standard.

Of the seven studies included for review, only one addressed the use of PET in detecting lymph node metastases (Wagner *et al.* 1999). This "prospective blinded" study enrolled 74 patients, 70 of whom had assessable cutaneous lesions > 1 mm in depth. PET was positive in only three of the 18 patients with positive SNB, corresponding to a sensitivity of 17%. Although specificity was 96%, PET failed to capture 83% of patients with positive SNB. This is an unacceptable number of patients to forgo or delay necessary lymph node dissection.

Application of working diagnostic guidelines:

- *Is this study of FDG PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

The Wagner study has no obvious sources of bias, although it was not clear whether consecutive patients were recruited in an effort to minimize selection bias.

- *What is the potential impact of FDG PET accuracy upon health outcomes?*

Based on the apparent lack of evidence presented above, coupled with the further lack of patient management and outcomes data, PET cannot replace SNB as a safe and less invasive method of detecting lymph node metastases.

Recommendation:	Coverage is not supported for using PET to evaluate regional lymph nodes.
Rationale:	It is clear that strongly negative studies should play an important role in providing requisite caveats. In this instance, where there is a lack of overwhelming evidence to the contrary, the Wagner study should be used to preclude PET coverage for regional lymph node evaluation.

Detecting extranodal metastasis at initial staging or during follow-up after treatment

As noted above, new coverage deliberations only apply to initial or primary, pre-treatment staging, given HCFA's reimbursement in monitoring for recurrent melanoma.

This evaluation focuses upon the addition of PET to conventional imaging (CI) studies and whether PET offers benefit to clinical decision-making. This obviously would allow for more appropriate, directed therapy if PET is more accurate (than CI) with respect to disease quantification and localization. Conversely, if PET either under-or overestimates disease, these patients will be inadvertently mistreated. The potential impact of this new technology depends upon the extent of discordance between conventional imaging and PET imaging. When both agree regarding either localized or metastatic disease, the management will not presumably change, but when there is discordance, the patient is at risk for harm. When PET falsely underestimates the extent of extranodal disease, patients receive less than optimal therapy. Conversely, when PET overestimates extranodal disease, patients may receive unnecessary therapy and are exposed to greater treatment morbidity.

The ideal study would prospectively categorize patients according to a reference standard stage of disease and compare the accuracy of identifying the stage with conventional imaging alone versus conventional imaging with PET. There were no studies available which were designed in this fashion. As an alternative, this review sought evidence which compared the diagnostic performance of PET and conventional imaging, whereby PET could demonstrate greater clinical utility if it was found to:

- Show better diagnostic performance;
- Be more often correct when discordant results are obtained;
- Accurately upstage or downstage patients, and
- Influence patient management NCDs.

The following study selection criteria were used in the Blue Cross/Blue Shield TEC Assessment:

- Published or accepted for publication as a full article in a peer-reviewed journal;
- At least 10 patients;
- Patient sample homogeneous with respect to type of primary cancer (i.e., studies excluded if there were either patients with various tumor types or if there was a mixture of primary and metastatic lesions);
- Performed tomographic rather than planar imaging with FDG as the radiotracer, and
- Correlation of PET findings with data from an appropriate reference standard, for at least some of the patients in the standard.

There were fifteen studies that met the selection criterion for melanoma. The most useful three include Rinne *et al.* (1998, n=100), Holder *et al.* (1998, n=76), and Valk *et al.* (1996, n=35). However, of these three, Rinne is most pertinent to the current question as it evaluated a subset of 52 patients who presented for initial staging. In this subgroup, the sensitivity of PET was 100% and the specificity was 94%, whereas conventional diagnostics did not identify any of the nine lymph node metastases (sensitivity = 0%) and also demonstrated a lower specificity (80%).

Application of working diagnostic guidelines:

- *Is this study of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

The Rinne study has no obvious sources of bias, although it was not clear whether consecutive patients were recruited in an effort to minimize selection bias.

- *What is the potential impact of PET accuracy upon health outcomes?*

There appears to be promising data, via Rinne *et al*., to support the use of PET when added to conventional imaging for the detection of metastases in melanoma patients, even though specific outcomes data is unavailable. However, it is likely that the expected clinical impact will be limited. In patients where there is a concordant result (with an expected majority of cases), there will be no significant change in management. The true impact will likely be realized when PET detects lesions missed by conventional imaging since patients will receive necessary treatment in a timely fashion without the delay from underdiagnosis.

Recommendation:	HCFA should add coverage for evaluating metastatic lesions during initial (primary) staging of malignant melanoma, in addition to its current coverage for recurrent melanoma.
Rationale:	The Rinne (1998) data has provided the evidentiary basis for broadening the scope of coverage for melanoma staging and recurrence.

Head and Neck Cancers (excluding malignancies of the central nervous system and thyroid)

Background

Cancer of the head and neck, excluding the central nervous system (CNS) and thyroid encompasses a diverse set of malignancies of which the majority are squamous cell carcinomas. These malignancies present at various sites, often arising in the oral cavity (lip, 45%; tongue, 16%; floor of the mouth, 12%; and buccal mucosa, 10%), and in various stages. The neck is a likely region metastatic spread of disease. Each of these sites has its own initial treatment protocol. Three clinical questions are addressed below:

- Identification of an unknown primary which has been detected as a metastasis in the neck;
- Initial staging of cervical lymph node metastases, and
- The detection of residual or recurrent disease following initial treatment.

Identification of an Unknown Primary Tumor

Patients may present with metastases to cervical lymph nodes but conventional forms of diagnostic imaging fail to identify the primary tumor. This leaves two options: Either neck dissection or radiation of both sides of the neck with random biopsies. PET scanning attempts to reveal the site of primary tumor to prevent the adverse effects of random biopsies or unneeded radiation. Beneficial outcomes might occur if PET accurately detects the primary site, and negates the need for biopsy or radiation of non-cancerous sites. Conversely, adverse outcomes occur when PET inaccurately identifies the site of primary cancer, thus permitting cancer to spread untreated throughout the body. If PET fails to identify a primary tumor site, the patient would be managed as having an unknown primary tumor.

PET could demonstrate greater clinical utility based upon its ability to accurately identify the site of a primary tumor. In assessing how often PET can identify a primary tumor, it is more useful to discuss the true positive rate. The true positive rate indicates how often PET accurately identifies the primary tumor among all patients tested. This allows the patient to forego radical neck dissection and/or diffuse radiation with random biopsies and the attendant morbidity associated with those treatments.

The recent Blue Cross/Blue Shield TEC assessment used the following criteria in selecting studies related to PET's use in locating unknown primary tumors of the neck:

- Published or accepted for publication as a full article in a peer-reviewed journal;
- At least 10 patients;
- Patient sample homogeneous with respect to type of primary cancer (i.e., studies excluded if there were either patients with various tumor types or if there was a mixture of primary and metastatic lesions);
- Performed tomographic rather than planar imaging with FDG as the radiotracer, and
- Correlation of PET findings with data from an appropriate reference standard, for at least some of the patients in the standard.

Of the eight studies addressing this issue, four were selected for review. Of the eight, six were prospective and two retrospective, while only one was blinded, two unblinded and five were unclear with regard to blinding. The primary distinction for inclusion focused upon the study's ability to consistently specify whether other imaging modalities were initially negative, thus more directly enabling determination of PET's incremental benefit. That was the case in the four studies selected. The pooled true positive rate (true positives/total number of patients) for all eight studies (n=138) was 32%. In the four studies (n=76) where patients had negative findings on both clinical examination and conventional imaging, the pooled true positive rate was 30%, while in those studies excluded (those not specifying whether other tests were initially negative), the pooled true positive rate for PET was 34%.

Application of working diagnostic guidelines:

- *Are the studies of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

The four studies selected did provide reasonable diagnostic performance data, and they are designed to extract reliable data. There are some limitations to the studies reviewed, most obviously the small sample sizes which can make sensitivity and specificity calculations unreliable.

- *What is the potential impact of PET accuracy upon health outcomes?*

The overall utility of PET appears positive. In cases where conventional imaging fail to find a primary and PET fails to find a primary, there is no change in management. If conventional imaging fails to find a primary and PET does find a likely site, but the biopsy fails to confirm this as primary, there is no change in management. The benefit of PET is where the PET identifies a primary that is confirmed by biopsy, and this leads to an initiation of directed tumor management. This scenario is statistically the least common, but may decrease the morbidity associated with unnecessary radiation and/or surgery. Unfortunately, these studies fail to show long-term survival for such patients. Therefore although there is a potential to demonstrate changes in management, it remains to be seen if this translates into real changes in outcomes.

Recommendation:	There is evidence to cover the use of FDG PET in the identification of unknown primary tumors with metastatic presentation in the neck.
Rationale:	Although the pooled studies demonstrate a relatively low true positive rate (30%), it is important to note that this rate represents the added diagnostic benefit of PET since the conventional work -up has already been noted to be negative. Thus, it is reasonable to support coverage if PET might be of assistance in nearly one-third of patients where diagnosis might otherwise have failed.

Initial Staging of Cervical Lymph Node Metastases

The decision to perform either neck dissection or irradiation is dependent upon the proper delineation of lymph node involvement by primary tumor. By first performing conventional imaging (CI), followed by PET, there are a few different possibilities:

- CI and PET are concordant such that treatment can be initiated which is suitable for that particular stage of cancer, or
- CI and PET are discordant such that, in turn, either:
    - PET downstages the disease (to lymph node negative) and less intensive therapy can be initiated (hence avoiding the adverse effects of this unnecessary therapy),
      or
    - PET upstages the disease (to lymph node positive), and more appropriate, intensive therapy can be initiated.

It should first be noted that a small (n = 19) prospective, blinded study by Wong *et al.* 1996 demonstrated the following:

- CT alone classified stage correctly in 69% of patients;
- CT first, then PET, classified stage correctly in 92% of patients;
- MRI alone with 40% correct staging, and
- MRI first, then PET, with 100% correct staging.

Seventeen studies considered by the Blue Cross/Blue Shield TEC Assessment reported improved head-to-head pooled sensitivities and specificities for either PET vs. CT or PET vs. MRI. This trend was consistent, regardless of whether the unit of analysis was number of neck sides, number of patients or number of lesions.

Application of working diagnostic guidelines:

- *Are the studies of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

In addition to the above inclusion criteria which may permit bias due to small sample sizes and a less than full comparison against a fixed reference standard, 11/17 studies have unclear blinding and 10/17 do not specify the desired recruitment of consecutive patients. Overall, however, there is a consistency in the finding of useful diagnostic information resulting from PET use.

- *What is the potential impact of PET accuracy upon health outcomes?*

Given the relatively stronger diagnostic performance illustrated above with respect to PET, coupled with the Wong study, it was inferred that more informed clinical decision making through the use of PET results might lead to improved health outcomes. However, no direct outcomes data were provided.

Recommendation:	FDG PET should be covered for the initial staging of cervical lymph nodes involved in metastatic disease.
Rationale:	The Wong study provided a small data set, but had a relatively strong design and demonstrated the benefit of PET. Additional confirmatory data was provided in other, less rigorous diagnostic trials.

The Detection of Residual or Recurrent Disease

Patients who have undergone surgery or radiation therapy often present with resultant tissue changes, such as scarring and fibrosis. This makes the identification of residual or recurrent tumor quite difficult via clinical examination, CI and even biopsy itself (on account of sampling discrepancies between biopsy sites themselves).

The TEC Assessment's causal chain logic in modeling this dilemma is as follows:

- Complete response with CI, then perform PET:
  - CI and PET are concordant: New treatment not needed
  - CI and PET are discordant: Can thus avoid delay in treating disease
- Recurrent/residual disease with CI, then perform PET:
  - CI and PET concordant: Confirmation of need to treat disease
  - CI and PET discordant: Can avoid adverse effects of unneeded treatment.

Using the above search criteria, 11 articles were determined to address the comparison of PET and CI modalities. However, before reviewing this list of articles, it should be noted that a small (n = 11) prospective, blinded study (also see above Wong *et al.* 1996) demonstrated the following:

- CT alone classified stage correctly in 88% of patients;
- CT first, then PET, classified stage correctly in 88% of patients;
- MRI alone with 50% (1/2 patients) correct staging, and
- MRI first, then PET, with 100% (both patients) correct staging.

The performance data for the 11 studies are sorted into three groupings based upon their pattern of findings. Six studies (Lowe *et al*. 2000, Wong *et al*. 1997, Anzai *et al*. 1996, Farber *et al*. 1998, Rege *et al*. 1994, Kao *et al*. 1999) demonstrated an overall relative superior sensitivity/specificity performance of PET, as compared with CT and/or MRI and physical examination (Lowe study only). An additional four studies provided neutral or mixed results (Hanasono *et al*. 1999, Manolidis *et al*. 1998, Nowak *et al*. 1999, Greven *et al*. 1997). Finally, a study by Paulus *et al.* 1998 reported overall less favorable diagnostic performance for PET relative to CT using data from local recurrences and lymph nodes.

Application of working diagnostic guidelines:

- *Are the studies of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

In addition to the above inclusion criteria which may permit bias due to small sample sizes and a less than full comparison against a fixed reference standard, 6/11 studies have unclear blinding and only a single study in this group specifies the desired recruitment of consecutive patients. Even with some conflicting sensitivity and specificity data, however, some useful diagnostic information is presented to support PET.

- *What is the potential impact of PET accuracy upon health outcomes?*

Given the relatively stronger diagnostic performance illustrated above with respect to PET, coupled with the Wong study, one may infer that a more informed clinical decision would lead to improved health outcomes. For example, in a series of 29 patients studied by Valk *et al.* (1996), PET findings were shown to avoid inappropriate surgery in nine patients (31%). Although no specific outcomes data were provided, it appears that earlier initiation of further treatment is possible if PET can detect recurrent disease when conventional imaging is negative.

Recommendation:	FDG PET should be covered for the detection of recurrent/residual tumor in patients with head and neck cancer.
Rationale:	While the Wong study provided a small data set, it had a relatively strong design and demonstrated the benefit of PET. Additional confirmatory data was provided in other small, less rigorous diagnostic trials.

Note: Separate requests for coverage of PET use for central nervous system and thyroid malignancies were included in the package received. However, they did not contain sufficient evidence to reach positive coverage determinations. PET use for central nervous system and thyroid malignancies remain non-covered indications at this time.

Myocardial Viability in Determining Coronary Revascularization

Background

Identification of patients with partial loss of heart muscle movement or hibernating myocardium is important in selecting candidates with compromised ventricular function to help determine appropriateness for revascularization. Diagnostic tests must distinguish between dysfunctional, yet viable myocardial tissue and scar tissue, in order to affect management NCDs. The decision to perform revascularization is based on the probability that improved systolic function that can occur with viable myocardium. FDG PET likely detects tissue that will not respond well to revascularization when single photon emission computed tomography (SPECT) is positive and FDG PET is negative.

The Commonwealth report evaluated the incremental benefit of FDG PET when SPECT has been used. It evaluated usefulness by assessing when SPECT has had negative or positive results compared to the FDG PET's negative or positive results thus creating four possible circumstances.

- SPECT positive and FDG PET positive
- SPECT negative and FDG PET negative
- SPECT positive and FDG PET negative
- SPECT negative and FDG PET positive

In scenario one or two there is essentially no change in management since results are concordant. In scenario three the proposed benefit of FDG PET would be to demonstrate that this is scar tissue with low likelihood of successful revascularization. As such the patient should be spared a procedure and exposure to complications. Conversely, in scenario four the sensitivity of FDG PET versus SPECT is challenged. When SPECT is negative but FDG PET is positive FDG PET must demonstrate there is evidence that revascularization improves outcomes.

To compare results, outcomes after revascularization must include, at a minimum, a change in ventricular wall motion. To be of incremental benefit, FDG PET must have greater sensitivity than SPECT, resulting in a larger number of successful revascularizations and improved outcomes for patients from improved systolic function. However, incremental benefit would also be achieved if unnecessary surgery was avoided because results indicated that revascularization would not be successful for SPECT positive/FDG PET negative findings.

Studies of interest would include patients who underwent both FDG PET and SPECT for pre-revascularization evaluation. Further, the patient's must be assessed after revascularization in a standard and acceptable fashion. Within this group patients of greatest interest would be those with discordant FDG PET and SPECT results. Ideally, these studies would look at patient-centered outcomes, instead, most use two-dimensional echo or ejection fraction (neither is the gold standard).

Lastly, Studies of the above type should also:

- Provide a clear description of patient entry characteristics
- All consecutive patients fulfilling criterion should be entered into the study
- FDG PET and SPECT should be done blinded to each other
- All those in categories of interest should be revascularized and followed-up

Application of Working Diagnostic Guidelines:

Are the studies of FDG PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?

Thirty-three full text papers were reviewed in the Commonwealth report. Of these none was specifically designed as outlined above. However, useful data was provided and is sufficiently accurate and free from bias. Multiple studies demonstrated that both SPECT and FDG PET identified viable myocardium (myocardium which recovered well after revascularization). FDG PET further predicts an improvement of heart failure symptoms and reduction in mortality. However, the present data is limited. The studies use echocardiography instead of ventriculography for assessment. The outcomes focus on surrogates and not long-term patient-centered outcomes. Lastly, the studies are predominately observational and prevent strong confidence in interpretation of results.

What is the potential impact of FDG PET accuracy upon health outcomes?

Maddahi et al 1994 reports FDG PET superior to SPECT as follows:

	SPECT	FDG PET
Sensitivity %	86	90
Specificity %	47	74
Positive Predictive Value %	72	83
Negative Predictive Value %	70	84

When results of FDG PET and SPECT were concordant this provided no change in management. In the setting of negative SPECT and positive FDG PET there was insufficient data to determine that the change in patient management resulting from the use of FDG PET would result in improved health outcomes. In fact only two studies demonstrated a change in management if FDG PET was used in addition to SPECT when SPECT was negative. One of the studies did not meet the inclusion criteria of consecutive patients and the other provided no evidence that outcomes were improved.

The use of FDG PET is promising when SPECT is positive but a question remains with regard to revascularization on clinical grounds. Three papers indicated a follow-up on outcomes after revascularization had taken place for patients who had both tests. However, data was not fully presented in two of the studies and the third study (Soufer 1995) did not revascularize all patients in the FDG PET positive SPECT negative group. Only 13 of the 37 patients identified received revascularization. The criteria that might have been used to reduce the number of patients receiving revascularization to 13 was not noted in the Assessment. Of more significance in the same study, there was no apparent improvement in regional ejection fraction and only one reported improvement in regional wall motion out of the 7 patients that were FDG PET negative and SPECT positive. It is this population which benefits from scanning because patients are spared side-effects of unnecessary procedures.

Recommendation: The use of FDG PET is supported for use when SPECT is positive and clinical correlation casts doubt on this finding to further predict myocardium amenable to revascularization.

Rationale: The evidence in Soufer (1995) was adequate to demonstrate the benefit of FDG PET when SPECT is positive but other clinical data does not support the test result for the purpose of avoiding unnecessary surgery.

Refractory Seizures

Background

A seizure is a transient disturbance of cerebral function, caused by an abnormal neuronal discharge, whereas epilepsy is a group of disorders characterized by recurrent seizures. Seizures can result from either primary central nervous system dysfunction or as a result of underlying metabolic derangement/systemic disease. Idiopathic epilepsy affects 0.2-0.4% of the general population. Whereas generalized seizures are characterized by loss of consciousness, complex partial seizures (marked by impaired consciousness) are the type most often targeted for surgical management when medical therapy has failed.

There have been divergent findings with respect to the benefit of FDG PET scanning in patients with refractory epilepsy where there is inconclusive localization of a seizure focus using non-invasive methods. Whereas approximately 25% of patients with seizure disorders have intractable (or refractory) seizures, 12-25% of these patients, in turn, are candidates for surgery, having failed medical therapy. Noting improvement rates exceeding 80% for temporal lobe resectable foci, extratemporal surgery has been somewhat less successful.

Potential Role of PET in Pre-Surgical Evaluation of Refractory Seizures

The Commonwealth assessment postulated the following key question: *Does PET have any incremental effect over the usual pre-surgical evaluation conducted to identify and delineate the epileptogenic foci?*

Several non-invasive diagnostic parameters include brain imaging, clinical/physical examination, neuropsychological testing, and surface electroencephalogram (EEG) testing; however, inconclusive testing can warrant invasive monitoring such as EEG with depth and grid electrodes. Therefore, if PET can provide additional non-invasive confirmation of seizure focus localization, then more patients might avoid preoperative invasive EEG.

The Commonwealth report first notes discrepant recommendations from earlier assessments, and elucidates selected methodological shortfalls. For example, a health technology assessment from AHCPR in 1998 had opposite recommendations from a 1997 TEC assessment. As a means of rectifying such differences, the Commonwealth report defines its own current review objectives:

- To summarize studies reporting the diagnostic accuracy of FDG PET in the localization of epileptogenic foci in patients who have undergone pre-surgical evaluation;
- To summarize studies which report the incremental benefit of PET in patients with refractory epilepsy being considered for surgery when there is no focus with concordant results on usual structural imaging and EEG, and
- To report studies which evaluate the effect of PET on decision-making and health outcomes.

The study inclusion criteria were as follows, with only five studies having been selected for final evaluation:

- Patients with epilepsy refractory to medical treatment being considered for surgery;
- Full articles reported in English;
- Conducted in humans;
- Should have reported information on diagnostic accuracy (or have provided sufficient information for it to be calculated) or should have specifically addressed the incremental benefit of PET, and
- Should have provided an adequate definition what constituted a "positive test," or provided information on the effect of PET on management NCDs.

None of the five articles directly report out their accuracy data such that the key issue of PET substitutability for invasive EEG can be addressed using sensitivity and specificity. However, the article by Delbeke *et al.*1996 provides some 2x2 frequency table performance data, which can be used to support the use of PET. In a series of 38 consecutive pre-operative patients, PET alone had a 94% positive predictive value for predicting significant post-surgical improvement, and for 22 patients in which invasive EEG was performed, 19 (86%) showed concordant hypometabolic foci with PET. These results were fairly similar for non-invasive EEG in which 30/36 patients (83%) demonstrated such concordant localization.

Application of working diagnostic guidelines:

- *Is this study of PET accuracy sufficiently free of bias to permit conclusions about the accuracy of PET as a diagnostic imaging test?*

The Delbeke study demonstrated several strengths including the use of consecutive patients, adequate blinding, and actual surgical outcomes. However, one drawback is the presence of work-up bias since not all relevant diagnostic procedures were performed on all patients.

- *What is the potential impact of PET accuracy upon health outcomes?*

The above noted correlation of surgical improvement with pre-operative PET scanning enables the ability to quantify this potential impact of PET upon health outcomes.

Recommendation:	There is some evidence to suggest the diagnostic benefit of PET in the pre-surgical management of patients with refractory seizures.
Rationale:	There is reference in the literature (Engel *et al.* 1990) that some surgical patients have already "skipped" invasive EEG on account of prior localization using PET (and at least 2 other non-invasive tests of focal functional deficit). Coupling this reference with the Delbeke data, PET would appear to have a viable role in the pre-surgical evaluation of refractory seizures.

Data on different types of PET cameras

In general, there is little data comparing the sensitivity and specificity achieved using different types of PET scanning systems. Extensive discussions with nuclear medicine experts reveals that there is considerable agreement that the quality of images produced by different systems can be markedly different. It is also clear that there are some scanning systems that are FDA approved for PET, but produce visibly lower quality than high end systems, such as the dedicated full ring BGO scanners. Given the importance of the clinical NCDs being made based on PET results, the quality of image production is a significant concern. The August 2000 "Report of the Commonwealth Review of Positron Emission Tomography" notes these alternative imaging systems to be inferior in sensitivity, especially for the detection of lesions measuring less than 1 cm. The Commonwealth's comparative evaluation of PET scanners was detailed and included discussion of spatial resolution, energy resolution, detection efficiency (sensitivity), count rate performance, noise equivalent count (NEC) rate, sensitivity to out of field of view activity, axial field of view, plus attenuation correction and image reconstruction.

Publications by American investigators have stated that some approaches "have failed to detect a large fraction of cancers in the clinically relevant 1 to 3 cm range, depending on the specific camera, the specific location in the body, and the tumor uptake" and that presently "the knowledge base is most secure for the dedicated full-ring PET imaging scanners, which are optimized for imaging positron emitters." (Macfarlane *et al.* 1995, Shreve *et al.* 1998, Wahl 1999).

Additionally, the October 2000 Seminars in Nuclear Medicine was largely devoted to coincidence imaging and included a comprehensive review of "The Role of Hybrid Cameras in Oncology" (Delbeke and Sandler 2000). Table 1 in that article (See Appendix A) Vanderbilt Experience with 511-keV Imaging Using a Dualhead Gamma Camera – corroborated markedly decreased detection rates for malignant lesions scanned on dualhead coincidence (DHC) gamma cameras versus dedicated PET (ECAT 933/08/16; CTI/Siemens, Knoxville, TN). For lesions less than 1.5 cm, lesion detection ranged from only 25% (on 3/8 inch collimated SPECT) to 61% (on 5/8 inch DHC) as compared to dedicated PET. Delbeke and Sandler noted that other authors have independently confirmed "lesion detection rates of the same range using dedicated PET images as the standard of reference" (Landoni *et al.* 1999, Zimny *et al.* 1999), and that "the limited detection rate using DHC in patients with oncologic disease has also been reported by other investigators with a system developed by a different manufacturer (Vertex MCD; ADAC Laboratories, Milpitas, CA) using 5/8 inch crystals" (Shreve *et al.* 1997, Shreve *et al.* 1998).

NATIONAL MEDICARE COVERAGE POLICY DETERMINATIONS

HCFA has concluded that the evidence available on use of FDG PET is sufficient to support broad coverage for diagnosis, staging and restaging for six types of cancer, and for limited diagnostic use for 2 non-oncologic indications. Details of this expanded coverage for FDG PET are provided in the table below, and the limitations on this coverage are described following the table. We have determined that the currently available evidence does not support broad coverage for all of the proposed clinical indications listed in the July 10, 2000 request. Therefore, use of PET for all other indications will remain non-covered.

Basis of expanded coverage for FDG PET

HCFA has decided that coverage for use of FDG PET for a specific type of cancer is approved for all clinically appropriate indications when one or more specific clinical indications for that cancer have been adequately demonstrated in scientific studies. This means that the conclusion that FDG PET is reasonable and necessary for all clinically appropriate uses within a single cancer type will be extrapolated from one or more empirically demonstrated clinical uses.

This approach to coverage is derived from an understanding of the novel underlying molecular basis of PET imaging. Specifically, PET images are produced as a result of the abnormal glucose metabolism of most malignant tissue. The metabolic abnormality associated with a particular cancer type does not vary depending on the specific diagnostic purpose for which the test is being used. Therefore, it is reasonable to conclude that the data on test performance and clinical utility of FDG PET produced through study of one indication of a particular cancer provides some information about the test performance and clinical utility of FDG PET for other clinical applications within the same cancer. HCFA expects that additional empirical study and further clinical experience will clarify the specific clinical uses for which FDG PET is most beneficial, but has determined that the reasonable and necessary threshold for PET is satisfied by an adequate scientific demonstration of one or more specific clinical indications for a specific type of cancer.

This approach to making a reasonable and necessary determination cannot necessarily be extended to clinical indications other than cancer diagnosis, because it is specific to diagnostic modalities that target general underlying metabolic abnormalities that are associated with the malignancies in question. In addition, the clinical utility of PET is largely derived from the nature of the clinical context in which PET is most commonly considered; that is when a clinician needs to decide whether to provide or withhold a potentially effective but clearly toxic or risky therapeutic intervention. The clinical utility of the diagnostic information provided by PET may be considerably less in circumstances where available treatments are not particularly effective or are associated with low toxicity or risk of harm. Whether this framework for determinations of reasonable and necessary is appropriate for other types of technologies or clinical entities will need to be determined on a case by case basis.

Limited quality of available studies

While we have determined that the available evidence was adequate to significantly expand Medicare coverage, the quality of evidence from available empirical studies of FDG PET was not consistent with the state-of-the-art in evaluating diagnostic tests. The characteristics of high quality studies are well-known, and are briefly described in the body of this document. Many of the studies we reviewed had serious methodologic limitations, making it difficult to arrive at clear conclusions about the benefit of FDG PET. The poor quality of empirical data for many clinical indications is not an academic or technical issue. The main concern is that results from poorly designed studies can lead to incorrect clinical NCDs and poor quality of care for Medicare beneficiaries. Experts in nuclear medicine and clinical medicine rely on information about the sensitivity and specificity of FDG PET in order to determine how likely it is that a positive or negative test result is actually true or false. The decision to proceed with or defer an invasive diagnostic procedure, surgery or chemotherapy therefore depends on reliable information about the performance characteristics of the test. Flawed scientific studies evaluating FDG PET may lead to incorrect interpretations of the test results, and patients may not receive the most appropriate care or may be inadvertently harmed.

Clinical experience and intuition alone are insufficient to determine the likelihood that a particular test result is true or false. That is the role of properly designed, objective empirical studies. Higher quality studies will inform higher quality clinical NCDs, leading to better health outcomes for patients. Poor quality studies may support incorrect NCDs that lead to patient harm. As additional studies of higher quality become available, it will be possible to reconsider this national coverage decision on FDG PET and make any revisions necessary to reflect the advancing state of knowledge about this technology.

Coverage is limited to selected high performance PET scanners only

The majority of the evidence submitted to HCFA and available in the scientific literature regarding the diagnostic performance of PET was derived from use of dedicated full ring bismuth germanate (BGO) PET scanners. As noted above in the last portion of the review of scientific evidence, available studies suggest that some other types of scanners may not perform as well as the full ring scanners, and may miss clinically important malignant lesions. Coverage for FDG PET is limited to use of dedicated full-ring PET scanners utilizing BGO, sodium iodide (NaI), or new crystal detector technologies that are equal or superior in performance. Also covered will be partial ring systems using BGO, partial ring NaI scanners with at least a 1" thick crystal, and scanners with new crystal detector technologies that are equal or superior in performance. Medicare will not cover any other scanning systems for performing PET, including gamma cameras modified for either non-coincidence or coincidence imaging. For those indications previously covered, PET scanners approved or cleared for marketing by the FDA remain covered.

HCFA is also aware that technology in this area is changing rapidly, and we are anxious to review any available data comparing the image quality, resolution and sensitivity of newer PET scanners to the data that currently exists relating to the high performance full ring PET scanners. A new coverage request containing comparative performance data will be required for HCFA to cover PET studies performed with scanners not listed in this paragraph.

Summary Table of New Medicare Coverage Policy for FDG PET

Clinical Condition	Coverage Decision (see limitations below)
Lung Cancer (non-small cell)	Diagnosis, staging and restaging
Esophageal Cancer	Diagnosis, staging and restaging
Colorectal Cancer	Diagnosis, staging and restaging
Lymphoma	Diagnosis, staging and restaging
Melanoma	Diagnosis, staging and restaging; Non-covered for evaluating regional nodes
Head and Neck Cancers (excluding CNS and thyroid)	Diagnosis, staging and restaging
Breast Cancer	Referred to MCAC Diagnostic Imaging Panel

Clinical Condition	Coverage Decision (see limitations below)
Myocardial Viability	Covered following inconclusive SPECT; Referred to MCAC Diagnostic Imaging Panel for review of possible additional uses
Refractory Seizures	Covered for pre-surgical evaluation
Alzheimer's Disease / Dementia	Referred to MCAC Diagnostic Imaging Panel
Remaining indications listed in the July 10, 2000 broad coverage request	Non-covered

For the three conditions being referred for consideration by the Medicare Coverage Advisory Committee (MCAC), HCFA will internally generate new requests for a national coverage decision.

Conditions and limitations for coverage:

- We do not believe that it is reasonable and necessary to cover specific clinical indications for which adequate scientific data demonstrate that PET does not provide medical benefit. When such evidence exists, use in these indications will be specifically excluded from coverage.
- For use in oncologic diagnosis: PET is covered in clinical situations in which the PET results may assist in avoiding an invasive diagnostic procedure, or in which the PET results may assist in determining the optimal location to perform an invasive diagnostic procedure. PET is not covered for other diagnostic uses, and is not covered for screening (testing of patients without specific symptoms).
- For staging and restaging: Coverage for PET is subject to 2 conditions: 1) the stage of the cancer remains in doubt after completion of a standard diagnostic workup, including conventional imaging, and 2) clinical management of the patient would differ depending on the stage of the cancer identified. Use of PET would also be considered reasonable and necessary if it could potentially replace one or more conventional imaging studies.
- We consider restaging to include both restaging in the setting of recurrence and restaging following completion of a therapeutic regimen or to assess whether a complete response has been achieved. Use of PET to monitor tumor response during the planned course of therapy (i.e. when no change in therapy is being contemplated) is not covered.

Prior to obtaining an FDG PET study, the physician ordering this imaging procedure will be required to document in the patient's chart the specific clinical question that will be answered by the imaging study. The ordering physician will thereby be certifying the medical necessity of the study according to the conditions described above. This documentation is necessary in order for HCFA to be able to reliably review the appropriateness of use of FDG PET under the expanded coverage described in this document. HCFA plans to conduct a review within the first year following the effective date of this new coverage, and will use the results of this review to determine whether there is any need for further review and to decide whether revisions to the coverage policy would be indicated.

Need for additional research

As noted above, the quality of studies that have been performed to evaluate FDG PET could be significantly improved. In all of the clinical conditions for which Medicare will now provide coverage, and for the remaining oncologic and other clinical uses, there is still a need for additional high quality clinical studies. HCFA is aware that there is limited public and private funding available for clinical research, particularly for studies that evaluate the clinical utility of promising technologies that emerge from basic research. For this reason, Medicare has recently implemented a policy for paying the routine costs for patients in clinical trials. The policy is aimed at increasing participation of Medicare patients in diagnostic and therapeutic trials, and well-designed evaluations of PET would be likely to qualify for coverage under this policy. For technologies of unique public health importance, HCFA will consider paying for the cost of experimental interventions in the context of clinical trials. This has been done in the past for several NIH-sponsored clinical trials that will provide critical evidence for developing HCFA coverage policy.

HCFA encourages the PET community to consult with experts in the evaluation of diagnostic technology in designing studies that will improve the empirical information available to clinicians and patients who use PET. HCFA staff is also available to meet with scientists and clinicians involved in the development of novel technologies in order to provide general advice on study design. We have initiated discussion with the National Cancer Institute to explore the possibility of collaborating with the PET community on these high priority studies, and look forward to continuing those discussions. More consistent conduct of these studies will be the most efficient way for Medicare to continue to expand coverage for novel beneficial technologies in a time frame that better matches the pace at which they are being developed.

Consideration of remaining indications

The current request for broad coverage received on July 10, 2000 is now considered closed by virtue of this coverage decision. Our review of all evidence submitted and additional evidence gathered supports the conclusion that the request for broad coverage is denied. Within that broad coverage request, we did find sufficient evidence to support coverage for the conditions described earlier in this document. The use of PET for clinical indications not addressed in this decision memo or previous Medicare coverage policies will remain non-covered. We encourage the requesters or others to submit new separate coverage requests for use of FDG PET in any additional clinical conditions that they believe would meet the coverage standards described in this document.

Appendix A

FDA Supporting Material form Submission

Appendix B
Evaluation of Diagnostic Tests
Table II/Fourfold table demonstrating "blind" comparison with "gold standard"
(Source: CMA Journal 1981)

		Gold standard		
		Patient has the disease	Patient does *not have* the disease	
Test result (conclusion drawn from the results of the test)	Positive: Patient appears *to have* the disease	True Positive a	False Positive b	a + b
	Negative: Patient appears *to not have* the disease	c False Negative	d True Negative	c + d
		a + c	b + d	a + b + c + d

Stable properties:
a/(a + c) = sensitivity
d/(b + d) + specificity

Frequency-dependent properties:
a/(a + b) = positive predictive value*
d/(c + d) = negative predictive value
(a + d)/(a + b + c + d) = accuracy
(a + c)/(a + b + c + d) = prevalence

*Positive predictive value can be calculated other ways too. One of them uses Bayes' theorem:

$$\frac{(\text{prevalence})(\text{sensitivity})}{(\text{prevalence})(\text{sensitivity}) + (1 - \text{prevalence})(1 - \text{specificity})}$$

Appendix C
MCAC Proposed Guidelines for Evaluating Diagnostic Tests

When they are asked to evaluate diagnostic tests, panels can apply criteria that are similar to those used for other health interventions that come before the Medicare Coverage Advisory Committee. The panels will need to determine whether the evidence is adequate to conclude that the diagnostic test improves outcomes and, if the evidence is adequate, to classify the magnitude of the health benefit, when a test is used for a specific purpose.

When more than one application of the test is under consideration, the panels will need to evaluate each application. Although this document refers to diagnostic tests, it is important to recognize that tests have four principal uses in clinical settings, and that the comments in this document refer to all four uses.

Screening: screening refers to the use of a test to detect either asymptomatic disease or a predisposition to disease (i.e., a risk factor such as elevated blood pressure or high blood cholesterol). Typically, the pre-test probability of disease (i.e., the prevalence or probability of disease in the population to be screened) is very low in such individuals. The purpose of screening is either to take action to prevent disease by modifying a risk factor, or to detect and treat disease early. In both cases, screening is presumed to be advantageous because early treatment of disease, or modification of a risk factor, improves health outcomes.

Diagnosis: a test is used to make a diagnosis when symptoms, abnormalities on physical examination, or other evidence suggests but does not prove that a disease is present. Making a correct diagnosis improves health outcomes by leading to better clinical NCDs about further testing and/or treatment.

Staging: a test is used to stage a disease when the diagnosis is known but the extent of disease is not known. Staging is particularly important when stage of disease, as well as the diagnosis itself, influences management. For example, an early stage cancer might be treated surgically, while the same cancer at a more advanced stage might be treated with chemotherapy alone.

Monitoring: in a patient known to have a health condition, a test is used to monitor the disease course or the effect of therapy. A monitoring test helps to evaluate the success of treatment and the need for additional testing or treatment.

Although an effective diagnostic test reduces the morbidity and mortality of disease by guiding clinical NCDs, direct proof of effectiveness is usually unavailable. Few studies have directly measured the effects of a diagnostic or screening test on health outcomes (studies of occult blood testing for colon cancer represent one such exception). Typical studies that evaluate the effectiveness of diagnostic, screening, or monitoring tests focus either on technical characteristics (e.g., does a new radiographic test produce higher resolution images) or effects on accuracy (does it distinguish between patients with and without a disease better than another test).

An improvement in the technical performance of a test can lead to improved diagnostic accuracy. For example, a higher resolution imaging study is more likely to distinguish between normal and abnormal anatomic structures, since it is able to delineate both types of structures more clearly. It may seem self-evident that improved technical characteristics would routinely lead to greater test accuracy and clinical utility, but that is not always the case. Often the factor that limits the ability of a test to distinguish between diseased and non-diseased, or between a person at high risk for disease and a person at average risk, is not the technical performance of the test. Sometimes the indicator that we are trying to measure (e.g., the risk factor) is only imperfectly correlated with the health condition, and improved measurement of the indicator will not lead to greater accuracy. Occasionally technical performance can improve in one respect but worsen in another; for example, MRI scans have higher resolution than most CT scans. Thus MRI scans were initially believed to be superior to CT scans for most indications. However, because CT scans are better able to distinguish certain tissue types, they proved to be better at detecting some abnormalities than the higher-resolution MRI scans. Thus improvements in aspects of technical performance are not sufficient to establish improved diagnostic accuracy.

When good quality studies directly measure how the use of a diagnostic test affects health outcomes, the panel can easily determine that the evidence is adequate and draw conclusions about the magnitude of the health benefits. But when the best studies only measure the accuracy of the test itself, the panels will have to determine whether the evidence is adequate to conclude that the test improves the accuracy of diagnosis or staging of disease *and* that the improvement in accuracy leads to better health outcomes.

We suggest that panels evaluating diagnostic test answer the following question:

Is the evidence adequate to conclude that the use of the diagnostic test leads to a clinically significant improvement in health outcomes?

If *direct* evidence linking the use of the test to health outcomes is not available, the panels should answer the following questions, which collectively determine whether there is convincing *indirect* evidence that the test will lead to better health outcomes:

Question 1: ***Is the evidence adequate to determine that the use of the test provides more accurate diagnostic information?***

The definition of "more accurate" is crucial. The standard measures of accuracy are **sensitivity** (probability of a positive test result in a patient with a disease or risk factor or other health condition) and **specificity** (the probability of a negative test result in a patient who does not have the disease). Ideally a new test would increase *both* sensitivity and specificity. Often that is not the case. A test that has a higher sensitivity is not unambiguously more accurate than an alternative test unless its specificity is at least as great. For most diagnostic tests, a change in the definition of an abnormal result will change the sensitivity, but improved sensitivity is obtained at the cost of worsened specificity, and vice versa. For example, if the diagnosis of diabetes is made on the basis of a fasting blood sugar, the use of a lower blood sugar level to define diabetes results in greater sensitivity and lowered specificity when compared to a diagnostic threshold at a higher blood glucose level. By choosing a different threshold, it is possible to change sensitivity without changing the test. Thus, if only sensitivity (or specificity) were considered, the same test might appear more accurate solely because the definition of an abnormal test result was changed.

The foregoing discussion leads to the following definition of "more accurate:" A more accurate test is not only more sensitive (or specific); it *has a higher sensitivity for a given level of specificity* when compared to another test. At a minimum, then, to conclude that one test is more accurate than another, its sensitivity (or specificity) is must be higher while its specificity (or sensitivity) is the same or better than the alternative test or diagnostic strategy.[1]

In deciding whether one test is more accurate than a second, established test, the panels will find the following steps helpful.

Step 1: Evaluate the quality of studies of test performance

The panel should first address the quality of the studies that are used to determine test accuracy. In assessing the quality of studies, panels might first consider the characteristics of an "ideal" study of test accuracy and compare the existing studies to the ideal. "Ideal" and "typical" studies of a screening, diagnostic, or monitoring test differ in these ways:

Ideal study	Usual study	Effect of Usual Study
The study subjects are consecutive patients seen in a typical clinical setting with a chief complaint.	Subjects selected because they had the diagnostic gold standard.	Overestimates sensitivity and underestimates specificity
All patients who get the index test also get the reference test	Patients with negative results on the index test often don't get the diagnostic gold standard	Overestimates sensitivity and underestimates specificity
The person who interprets the index test is blinded to all other information		Overestimates sensitivity and specificity.

Ideal study	Usual study	Effect of Usual Study
	The person who interprets the index knows the clinical history and the results of the diagnostic gold standard.	
The person who interprets the reference test is blinded to all other information	The person who interprets the diagnostic gold standard knows the clinical history and the results of the index test.	Overestimates sensitivity and specificity.
The reference test is a valid measure of the disease state	The diagnostic gold standard imperfectly measures the disease state.	The measured test performance could either be worse or better than the true performance.

*The **reference test** is a test that is considered the "gold standard," i.e., a test that is used to define the disease. Tests commonly used as reference tests are coronary angiography, for coronary artery disease, and histopathology, for cancer. Reference test can be interpreted more broadly to mean any method that is considered the definite basis for determining whether a disease or risk factor is truly present.

The panels will need to decide whether the results of studies that fall short of the ideal are likely to be due to bias, or whether their limitations are sufficiently minor that it is possible to draw conclusions about the accuracy of the test.

Step 2: Evaluate the possibility that the two tests are complementary

The sensitivity and specificity of a new test can be the same as – or even worse than – the sensitivity and specificity of an established comparison test, yet still provide valuable information. It can add value if it provides complementary information. In this circumstance, a combination of the two tests leads to more accurate distinction between patients with and without the disease (or risk factor) than either test individually. The information is likely to be complementary if the other test or tests detect other features of the disease (for example, one test measures a physiological phenomenon while the other is an imaging test that detects structural abnormalities). A direct comparison between strategies using the two tests and those using only the standard test can be made by studying patients who receive both tests as well as the reference test (or any direct measure of whether disease is actually present). The appendix describes how such a study can be used to determine whether the combined testing strategy improves the accuracy of diagnosis.

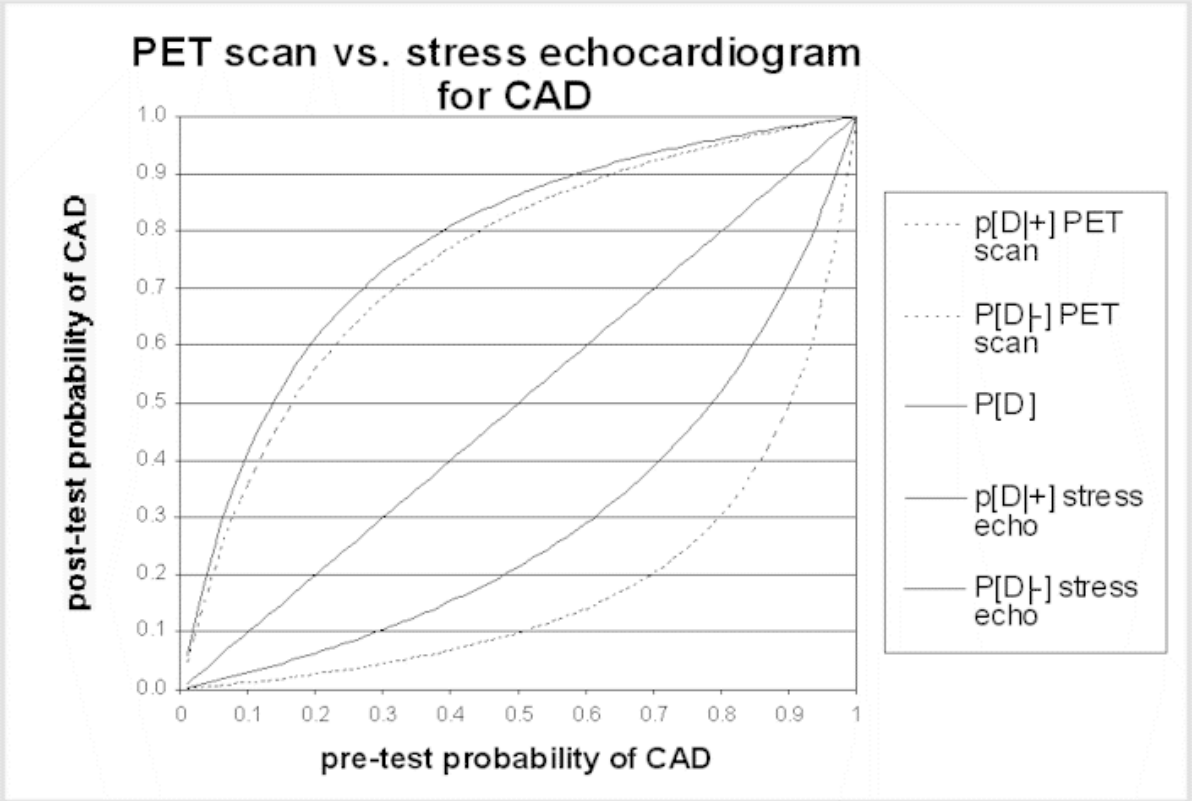
Question 2: ***If the test improves accuracy, is the evidence adequate to conclude that the improved accuracy will lead to better health outcomes?***

To determine whether a difference in test accuracy would lead to important improvements in health outcomes, the panels may find the following steps helpful.

Step 1: Calculate the post-test probability of disease

The purpose of testing is to reduce uncertainty about the presence of a disease or risk factor, or about the extent of a previously diagnosed disease. The pre-test probability of disease is the probability of disease before the test has been performed, based upon history, physical examination, and preliminary diagnostic tests. The pre-test probability is often used interchangeably with the term "disease prevalence," but the two terms are only equivalent when prevalence and pre-test probability are based on the same population (i.e, adjusted for history and other information).

The post-test probability is the probability of disease after learning the test results. A test result should only change patient management if it changes the probability of disease. Bayes' theorem is the formal approach used to calculate the post-test probability. Application of Bayes' theorem in this context requires the sensitivity and specificity of the test and the pre-test probability of disease. Generally, tests alter probability the most (i.e., in comparison to the pre-test probability) when the pre-test probability is intermediate (i.e., not near a probability of either 0 or 1). Conversely, tests alter probability the least when the pre-test probability is close to zero or close to 1.0. If the patient's symptoms, abnormalities on physical examination, and other evidence strongly suggest that the patient has the disease in question (i.e., the pre-test probability of disease is high), unless a test is extremely sensitive the patient is likely to have the disease even if the test result is negative, and should be managed accordingly. Similarly, if the pre-test risk of disease is very low, the probability of disease in a patient with a positive test result remains very low, unless the test is extremely specific (i.e., rarely produces false-positive results). The accompanying graph of post-test probability for two tests illustrates this point. Panels may find these graphs helpful in interpreting the possible impact of a difference in test performance.



The same principles apply to the use of testing to stage disease or to monitor the effect of treatment. In these situations, the uncertainty is not about the diagnosis, but the test is needed to reduce uncertainty about the current status of the disease. Learning more about stage or response to treatment is important insofar as it will influence management options – for example, disease progression while on one treatment will often lead to a change in therapies, or cessation of a potentially toxic therapy. A false-negative staging test result (i.e., one that implies the disease is more limited than it really is) may lead to treatment that is both ineffective and harmful. In some situations, a false-positive staging test result can have even more harmful consequences; the physician could withhold potentially curative treatment if he or she interprets the staging test as indicating that cure is not possible, dooming a patient to die of a disease that could have been treated effectively.

Step 2: Evaluate the potential impact on management when tests differ in the post-test probability:

In the absence of direct evidence of the effects of a test on health outcomes, it will sometimes be possible to conclude with great confidence that improved accuracy will lead to better outcomes. This is particularly likely to be true when the treatment or management strategy is effective for patients with the disease, but poses risks or discomfort that would not be acceptable when administered to patients who do not have the disease. Then, improved accuracy leads to effective treatment for more people who truly have the disease, and helps avoid unnecessary treatment in people who would not benefit from it. Thus, although the evidence that diagnostic tests for cancer and for heart disease alter health outcomes is largely indirect, it is also compelling. For these categories of disease, there is often strong evidence that treatments with significant adverse consequences are effective when used appropriately. Panels will need to judge whether the test leads to better patient management by increasing the rate at which patients with disease receive appropriate treatment and the rate at which patients who do not have the disease avoid unnecessary treatment.

If management changes, the improvement in health outcomes should be large enough that the panel believes it is clinically significant. A small increase in accuracy can lead to substantial improvements in health outcomes if treatment is highly effective. Improved accuracy is of little consequence, however, if treatment is either ineffective, so there is little benefit to patients with the disease, or very safe, so there is little harm to patients without the disease. Then improved accuracy is unlikely to lead to improved health outcomes or even to influence clinical NCDs.

Under exceptional circumstances, prognostic information, even if it did not affect a treatment decision, could be considered to improve health outcomes. The panel should be alert for circumstances in which patients would be likely to value the prognostic information enough to significantly alter their well-being.

Summary

The recommended approach for evaluating diagnostic tests is as follows:

- Review, when available, high quality studies that provide *direct* evidence that test results improve health outcomes.

- If there is no high quality *direct* evidence, evaluate the *indirect* evidence as follows:

Decide whether studies of test accuracy are sufficiently free of bias to permit conclusions about the accuracy of the test under consideration, in comparison either to another test or another screening, diagnostic, or staging strategy

Evaluate the potential impact of improved accuracy (or complementary information) on health outcomes. Evaluating the effect of test accuracy on post-test probability is one part of this step. The other part is deciding whether the change in patient management that results from the test will improve health outcomes. Improved outcomes are likely to occur when the management strategy is effective in patients with the disease and does not benefit those without the disease. A test can also improve health outcomes when the treatment poses significant risk, so that it is very important to avoid unnecessary treatment.

---

[1]The more technical expression of this condition is that a more accurate test is one whose receiver operating characteristic (ROC) curve is above and to the left of the ROC curve for the alternative test.

APPENDIX: THE COMPLEMENTARY VALUE OF COMBINED TESTING

To test the hypothesis that two tests are complementary, several approaches are possible. The best way is a study in which a series of patients receive both tests as well as the reference test. The analysis compares the sensitivity of the second test in two groups of patients: those with a negative result on the first test and those with a positive result, as shown in the table.

	Test 1 results positive		**Test 1 results negative**	
Test 2 results	Reference standard positive	Reference standard negative	Reference standard positive	Reference standard negative
Positive	A		A'	
Negative	B		B'	
Totals	A+B		A'+B'	

If the sensitivity of Test 2 when test 1 is negative (A'/[A'+B']) is greater than zero, Test 2 is able to detect patients that Test 1 cannot, and the two tests are complementary. If, on the other hand, the sensitivity of Test 2 is zero when Test 1 is negative, Test 2 is unable to detect patients that Test 1 would miss, and it is of minimal additional value.

Many studies of two tests do not provide the information in this table. However, the studies may still provide useful data that reflect what is in the table. The best way to think about using two tests is to consider them as a sequence of tests, in which the post-test probability after the first test becomes the pre-test probability for the second test. Suppose that the test under consideration is the second test in the sequence. It would add information when compared to the established test alone under two circumstances:

- The first test in the sequence is positive, and the post-test probability after a positive result on the second test in the sequence is greater than the post-test probability after the first test.
- The first test in the sequence is negative, and the post-test probability after a negative result on the second test in the sequence is lower than the post-test probability after the first test.

Arguments that consist largely of inductive reasoning (based upon a different physiological basis for Test 2) are much weaker than empirical evidence.

Appendices [PDF, 325KB]
Back to Top

Bibliography

Abdel-Nabi H, Doerr RJ, Lamonica DM, *et al.*, Staging of primary colorectal carcinomas with fluorine-18 fluorodeoxyglucose whole-body PET: correlation with histopathologic and CT findings. *Radiology* 1998;206:755-760.

Anzai Y, Carroll WR, Quint DJ, *et al.,* Recurrence of head and neck cancer after surgery or irradiation: prospective comparison of 2-deoxy-2-[F-18]fluoro-D-glucose PET and MR imaging diagnosis. *Radiology* 1996;200:135-141.

Beanlands RS, deKemp RA, Smith S, Johansen H, Ruddy TD. F-18-fluorodeoxyglucose PET imaging alters clinical decision making in patients with impaired ventricular function. *American Journal of Cardiology* 1997;79:1092-1095.

Block MI, Patterson GA, Sundaresan RS, *et al.,* Improvement in staging of esophageal cancer with the addition of positron emission tomography. *Ann Thorac Surg* 1997; 64:770-777.

Bury T, Corhay JL, Duysinx B, *et al.,* Value of FDG PET in detecting residual or recurrent nonsmall cell lung cancer. *European Respiratory Journal* 1999; 14:1376-1380.

Chee MW, Morris HH 3rd, Antar MA, *et al*., Presurgical evaluation of temporal lobe epilepsy using interictal temporal spikes and positron emission tomography. *Arch Neurol* 1993;50:45-48.

Choi JY, Lee KH, Shim YM, *et al.,* Improved Detection of Individual Nodal Involvement in Squamous Cell Carcinoma of the Esophagus by FDG PET, *J of Nucl Med* 2000;41:5:808-815.

Delbeke D, Lawrence SK, Abou-Khalil B, *et al*., Postsurgical outcome of patients with uncontrolled complex partial seizures and temporal lobe hypometabolism on 18 FDG-positron emission tomography. *Investig Radiol* 1996;31:261-266.

Delbeke D, Sandler MP, The Role of Hybrid Cameras in Oncology. *Seminars in Nuclear Medicine* 2000;30:4:268-280.

Engel J Jr, Henry TR, Risinger MW, *et al*., Presurgical evaluation for partial epilepsy: relative contributions of chronic depth-electrode recordings versus FDG PET and scalp-sphenoidal ictal EEG. *Neurology* 1990;40:1670-1677.

Farber LA, Benard F, Machtay M, *et al*., Detection of recurrent head and neck squamous cell carcinomas after radiation therapy with 2-18F-fluoro-2-deoxy-D-glucose positron emission tomography. *Laryngoscope* 1999;109:970-975.

FDG Positron Emission Tomography in Colorectal Cancer, Blue Cross/Blue Shield TEC Assessment Program 2000;14:25.

FDG Positron Emission Tomography in Head and Neck, Blue Cross/Blue Shield TEC Assessment Program 2000; 15:4.

FDG Positron Emission Tomography in Lymphoma, Blue Cross/Blue Shield TEC Assessment Program 2000; 14:26.

FDG Positron Emission Tomography in Melanoma, TEC Blue Cross/Blue Shield TEC Assessment Program 2000;14:27.

Flamen, P, Leurt A, Van Custsem E, *et al.,* Utility of positron emission tomography for the staging of patients with potentially operable esophageal carcinoma. *J of Clin Oncol* 2000;18:18:3202-3210.

Flanagan, FL, Dehdashti F, Siegel, BA, *et al.,* Staging of esophageal cancer with ^{18}F-fluorodeoxyglucose positron emission tomography. *AJR* 1997;168:417-424.

Greven KM, Williams DW 3rd, Keyes JW, *et al.*, Can positron emission tomography distinguish tumor recurrence from irradiation sequelae in patients treated for larynx cancer? *Cancer J Sci Am* 1997;3:353-357.

Hanasono MM, Kunda LD, Segall GM, *et al.*, Uses and limitations of FDG positron emission tomography in patients with head and neck cancer. *Laryngoscope* 1999;109:880-885.

Haynes, RB, How to read clinical journals. *CMA Journal* 1981; 124:703-710.

Holder WD Jr, White RL Jr, Zuger JH, *et al.*, Effectiveness of positron emission tomography for the detection of melanoma metastases. *Ann Surg* 1998;227:764-769.

Kao C-H, ChangLai S-P, Chieng P-U, *et al.,* Detection of recurrent or persistent nasopharyngeal carcinomas after radiotherapy with 18-fluoro-2-deoxyglucose positron emission tomography and comparison with computed tomography. *J Clin Oncol* 1998;16:3550-3555.

Kole AC, Plukker JT, Nieweg OE, Vaalburg, W, Positron emission tomography for staging of oesophageal and gastroesophageal malignancy. *British J of Cancer* 1998;78:4:521-527.

Landoni C, Gianolli L, Lucignagni G, *et al.,* Comparison of dual-head coincidence PET versus ring PET in tumor patients. *J Nucl Med* 1999;40:1617-1622.

Lowe VJ, Boyd JH, Dunphy FR, *et al*., Surveillance for recurrent head and neck cancer using positron emission tomography. *J Clin Oncol* 2000;18:651-658.

Luketich JD, Friedman DM, Weigael TL, *et al.,* Evaluation of Distant Metastases in Esophageal Cancer: 100 Consecutive Positron Emission Tomography Scans. *Ann Thorac Surg* 1999;68:1133-1137.

Luketich JD, Schauer PR, Meltzer CC, *et al.,* Role of positron emission tomography in staging esophageal cancer. *Ann Thorac Surg* 1997;64:765-769.

Macfarlane DJ, Cotton L, Ackermann RJ, Minn H, Ficaro EP, Shreve PD, Wahl RL, Triple-head SPECT with 2-[fluorine-18] fluoro-2-deoxy-D-glucose (FDG): initial evaluation in oncology and comparison with FDG PET. *Radiology* 1995;194:425-429.

MacIntyre WJ, Go RT, King JL, *et al*., Clinical Outcome of Cardiac Patients with negative Thallium-201 SPECT and positive Rubidium-82-PET myocardial perfusion imaging. *Journal of Nuclear Medicine* 1993;38:742-748.

Maddahi J, Schelbert H, Brunken R, DiCarlin M. Role of thallium-201 and PET imaging in evaluation of myocardial viability and management of patietns with coronary artery disease and left ventricular dystfunction. *Journal of Nuclear Medicine* 1994;35:129-136.

Management Decision and Research center, Health Services Research and Development Service, Veterans Health Administration, Department of Veterans Affairs, Evaluating diagnostic tests: a guide to the literature. Technology Assessment Program Report No. 4, Boston: December, 1997.

Manolidis S, Donald PJ, Volk P, *et al*., The use of positron emission tomography scanning in occult and recurrent head and neck cancer. *Acta Otolaryngol Suppl (Stockh)* 1998:5341.

Nowak B, Di Martino E, Janicke S, *et al*., Diagnostic evaluation of malignant head and neck cancer by F-18-FDG PET compared to CT/MRI. *Nuklearmedizin* 1999;38:312-318.

Paulus P, Sambon A, Vivegnis D, *et al*., 18FDG PET for the assessment of primary head and neck tumors: clinical, computed tomography, and histopathological correlation in 38 patients. *Laryngoscope* 1998;108:1578-1583.

Que TH, Plukker JT, Vaalburg W, Staging of gastroesophageal cancer using C^{11} choline and F^{18} FDG. *Groningen University Hospital* 2000; Presentation: Clinical Tracer Comparisons.

Rege S, Maass A, Chaiken L, *et al*., Use of positron emission tomography with fluorodeoxyglucose in patients with extracranial head and neck cancers. *Cancer* 1994;73:3047-3058.

Report of the Commonwealth Review of Positron Emission Tomography, Health Access and Financing Division, Commonwealth Department of Health and Aged Care, 2000.

Rinne D, Baum RP, Hor G, *et al.*, Primary staging and follow-up of high risk melanoma patients with whole-body 18F-fluorodeoxyglucose positron emission tomography: results of a prospective study of 100 patients. *Cancer* 1998;82:1664-1671.

Schiepers C, Penninckx F, DeVadder N, *et al.*, Contribution of PET in the diagnosis of recurrent colorectal cancer: comparison with conventional imaging. *Eur J Surg Oncol* 1995;21:517-522.

Shreve PD, Anzai Y, Wahl RL, Pitfalls in oncologic diagnosis with FDG PET imaging: physiologic and benign variants. *Radiographics* 1999;19:61-77.

Shreve PD, Steventon RS, Deters E, *et al.*, FDG imaging of neoplasms using a dual head SPECT camera operated in coincidence mode. *Eur J Nucl Med* 1997;24:860.

Shreve PD, Steventon RS, Deters EC, Kison PV, Gross MD, Wahl RL, Oncologic diagnosis with 2-[fluorine-18]fluror-2-deoxy-D-glucose imaging: dual-head coincidence gamma camera versus positron emission tomographic scanner. *Radiology* 1998;207:431-437.

Silberstein EB, Prevalence of adverse reactions to positron emitting radiopharaceuticals in nuclear medicine. *J of Nucl Med* 1998; 39: 2190-2192.

Soufer R, Dey HM, Ng CK, Zaret BL, Comparison of sestamibi single-photon emission computed tomography with positron emission tomography for estimating left ventricular myocardial viability. *American Journal of Cardiology* 1995; 75:1214-1219.

Stumpe KD, Urbinelli M, Steinert HC, *et al*., Whole-body positron emission tomography using fluorodeoxyglucose for staging of lymphoma: effectiveness and comparison with computed tomography. *Eur J Nucl Med* 1998;25:721-728.

Systematic evaluation of material submitted to HCFA to support coverage of positron emission tomography, New England Medical center Evidence-Based Practice center, September 2000.

Valk PE, Abella-Columna E, Haseman MK, *et al.*, Whole-body PET imaging with [18F]fluorodeoxyglucose in management of recurrent colorectal cancer. *Arch Surg* 1999;134:503-511.

Valk PE, Pounds TR, Tesar RD, *et al.*, Cost-effectiveness of PET imaging in clinical oncology. *Nucl Med Biol* 1996;23:737-743.

Wagner JD, Schauwecker D, Hutchins G, *et al.*, Prospective study of fluorodeoxyglucose-positron emission tomography imaging of lymph node basins in melanoma patients undergoing sentinel node biopsy. *J Clin Oncol* 1999;17:1508-1515.

Wahl, RL, Positron Emission Tomography in Cancer Patient Management. *Scientific American* 1999; 5:4:205-206.

Wong WL, Chevretton EB, McGurk M, *et al*., A prospective study of PET-FDG imaging for the assessment of head and neck squamous cell carcinoma. *Clin Otolaryngol* 1997;22:209-214.

Wong W-L, Hussain K, Chevretton E, *et al*., Validation and clinical application of computer combined computed tomography and positron emission tomography with 2-[18F]fluoro-2-deoxy-D-glucose head and neck images. *Am J Surg* 1996;172:628-632.

Zimny M, Kaiser HJ, Cremerius U, *et al.,* F-18-FDG positron imaging in oncological patients: gamma camera coincidence detection versus dedicated PET. *Nuklearmedizin* 1999;38:108-114.

Zinzani PL, Magagnoli M, Chierichetti F, *et al.*, The role of positron emission tomography (PET) in the management of lymphoma patients. *Ann Oncol* 1999;10:1181-1184.

Back to Top