Diagnostic accuracy of 18F-fluorodeoxyglucose positron emission tomography in the follow-up of papillary or follicular thyroid cancer


Authors' objectives
To determine the diagnostic accuracy of positron emission tomography (PET) using 18F-fluorodeoxyglucose (FDG) in patients suspected of recurrent papillary or follicular thyroid carcinoma.

Searching
MEDLINE, EMBASE and Cancerlit were searched up to October 2000 for studies reported in any language. A specifically designed search strategy for diagnostic accuracy studies was combined with a specific search for PET, FDG and thyroid cancer. Further details of the searches were provided in the review. The Cochrane Library (Issue 2, 1999) was also searched. The references of relevant publications and reviews were searched for additional studies.

Study selection

Study designs of evaluations included in the review
Prospective or retrospective diagnostic accuracy studies with at least 10 participants were included. Reviews, case reports, editorials, letters and comments were excluded.

Specific interventions included in the review
Studies of PET using FDG were eligible for inclusion. Studies on FDG imaging using gamma cameras were excluded.

Reference standard test against which the new test was compared
No inclusion criteria relating to the reference standard were specified. However, the following reference standards were considered acceptable by the authors: histology or cytology; focal 131-I-uptake; pathognomic bone scan or magnetic resonance imaging for bone metastases; computed tomography or magnetic resonance imaging for brain metastases; and progression of radiologically documented lesions suspect for malignancy.

Participants included in the review
Studies of patients suspected of recurrent follicular and papillary thyroid cancer were included. Where reported, the mean age of the patients ranged from 41 to 67 years.

Outcomes assessed in the review
No inclusion criteria relating to the outcomes were specified. The outcomes reported in the review included sensitivity, specificity and positive and negative predictive values.

How were decisions on the relevance of primary studies made?
Two unblinded reviewers independently selected the studies, and any differences were resolved by consensus.

Assessment of study quality
The validity of the studies was assessed using the following criteria: reference standard; independence of interpretation; uniform application of reference test; comparison of different tests in a valid design; study design (prospective or retrospective); and missing data. Two unblinded reviewers independently assessed the methodological quality of the included studies, and any differences were resolved by consensus.

Data extraction
Two reviewers independently extracted the data. The data extracted on appropriate clinical setting and patient spectrum included the age distribution, gender distribution, number of patients, tumour type, tumour stage, and level of serum Tg measurements. The data relating to the description of the imaging procedure included patient preparation, timing and
duration of acquisition, FDG dose, attenuation correction, and scanned trajectory.

Data were also extracted on the proportion of patients with positive and negative PET scans. These were classified according to the defined set of criteria: true positive if confirmed by a valid reference standard and false positive if confirmed by histopathology. Patients with unconfirmed PET lesions and raised serum markers were classified as unclear. Patients with discrepancy of PET and reference standard and low serum markers were classified as false positive. Negative PET findings were classified as true negatives if confirmed by histopathology. Patients with congruent negative findings on PET and a valid reference standard, combined with a follow-up of 23 months, were also classified as true negative. Patients with discrepancy of PET and a valid reference test were classified as false negative. All other PET findings were classified as unclear.

Methods of synthesis
How were the studies combined?
The studies did not supply appropriate data and were too heterogeneous to allow a quantitative meta-analysis. The studies were analysed separately for three subgroups of patients: those with negative 131-I whole-body scintigraphy and raised serum markers; those with other clinical suspicion of relapse; and those with known neoplastic foci to complete the work-up.

How were differences between studies investigated?
Differences between the studies were not formally investigated, although they were discussed in the text.

Results of the review
Fourteen studies (n=402) were included: 7 were prospective, 5 retrospective, and for 2 the design was unclear.

Most studies lacked information on one or more quality items. In particular, only small groups of patients were submitted to valid reference tests; the results in patients with negative FDG PET were often not confirmed in a valid way; most interpretations of FDG PET and the reference test were probably not performed independently of each other; the selection of patients for the assessment by the reference test was often not independent of the FDG PET results; and in most studies there was no description regarding missing data.

Data on sensitivity and specificity were only available for 7 studies. The sensitivity ranged from 70% (with a specificity of 77%) to 95% (specificity not reported), while the specificity ranged from 77% (sensitivity 70%) to 100% (sensitivity 72%).

Authors' conclusions
The results seem to support the potential of FDG PET to identify and localise foci of recurrent cancer in patients with elevated serum Tg and negative iodine-131 scans. However, implementation of PET in a routine diagnostic algorithm requires additional evidence.

CRD commentary
This was a good review of the area. A reasonable literature search was conducted, although it is possible that it may have missed some important studies. The review answered a clear review question and the inclusion criteria were clearly reported. Details of the review process were adequate and suggested that appropriate steps had been taken to minimise bias. Details of the studies were tabulated. The synthesis used appears to have been appropriate, although it can be difficult to follow at times.

The authors' conclusions are supported by the results presented.

Implications of the review for practice and research
Practice: The authors state that 'the present evidence does not allow implementation of PET in a routine diagnostic
algorithm’.

Research: The authors did not state any implications for further research.

Bibliographic details

PubMedID
11502811

DOI
10.1210/jcem.86.8.7777

Original Paper URL
http://jcem.endojournals.org/cgi/content/full/86/8/3779

Indexing Status
Subject indexing assigned by NLM

MeSH
Adenocarcinoma, Follicular /diagnosis /radiography /radionuclide imaging; Brain Neoplasms /diagnosis /radiography /radionuclide imaging /secondary; Carcinoma, Papillary /diagnosis /radiography /radionuclide imaging; Databases as Topic; Fluorodeoxyglucose F18; Follow-Up Studies; Gamma Cameras; Humans; MEDLINE, Magnetic Resonance Imaging; Radiopharmaceuticals /therapeutic use; Recurrence; Reproducibility of Results; Thyroid Neoplasms /diagnosis /radiography /radionuclide imaging; Tomography, Emission-Computed; Tomography, X-Ray Computed

AccessionNumber
12001002001

Date bibliographic record published
31/01/2003

Date abstract record published
31/01/2003

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.