CRD summary
This review assessed the accuracy of 18F-fluorodeoxyglucose positron emission tomography (FDG-PET) in detecting lymph node metastases in head and neck squamous cell carcinoma and compared it with conventional tests (computed tomography, magnetic resonance imaging and ultrasound with fine-needle aspiration). The authors’ conclusion that all of these tests have similarly poor performance is likely to be reliable.

Authors' objectives
To assess the diagnostic accuracy of 18F-fluorodeoxyglucose positron emission tomography (FDG-PET) in detecting lymph node metastases in patients with head and neck squamous cell carcinoma (HNSCC) and to compare its performance with other tests: computed tomography (CT); magnetic resonance imaging (MRI); and ultrasound with fine-needle aspiration (USFNA).

Searching
MEDLINE was searched to July 2007. Search terms, based on the index test and the target condition, were reported. The bibliographies of retrieved articles were screened for additional studies. No language restrictions were applied.

Study selection
Studies that evaluated FDG-PET for detecting lymph node metastases at initial staging before surgery in patients with HNSCC were eligible for inclusion. Studies had to include at least five patients with HNSCC and patients with and without lymph node metastases as determined by histopathologic examination (reference standard). Studies with verification bias (those in which the reference standard of histopathologic examination was performed only in FDG-PET-positive patients) were excluded. Studies that used FDG-PET to assess post-treatment recurrence and studies in which patients received chemotherapy or radiotherapy before neck dissection were excluded. FDG-PET positivity was defined as qualitative for most included studies. All but two of the included studies assessed at least one comparator test in addition to FDG-PET. Where reported, the proportion of cN0 (clinically negative cervical lymph nodes) patients ranged from 0 to 100%.

The authors stated neither how studies were selected for the review nor how many reviewers performed the selection.

Assessment of study quality
The authors did not state that they assessed study validity, but studies with verification bias were excluded and study design (prospective or retrospective) and blinding were recorded.

Data extraction
The absolute numbers of true positive, false negative, false positive and true negative results were extracted for each test (FDG-PET, CT, MRI and USFNA). Investigators were contacted where data were missing. Data were also extracted on the clinical node staging of study populations: patients with clinically positive cervical lymph nodes (cN1, cN2 and cN3), patients with clinically negative cervical lymph nodes (cN0) and mixed populations.

Two reviewers independently extracted data. Differences were resolved by discussion or consultation with a third reviewer.

Methods of synthesis
Diagnostic performance data were combined using a hierarchical regression model to generate summary receiver operating characteristic (SROC) curves and summary estimates of per patient sensitivity and specificity, with 95% confidence intervals (CIs), for FDG-PET. The summary estimates of sensitivity and specificity derived from meta-
analyses were used to calculate positive and negative likelihood ratios (LRs) and negative predictive values (NPVs) for a range of population prevalences of node metastases. The hierarchical regression model was used to compare the diagnostic performance of FDG-PET with that of conventional tests (CT, MRI, CT/MRI and USFNA) individually or as a group; for group data a positive result was defined as a positive result from at least one conventional test.

Subgroup analyses were performed for clinical N stage (cN positive, cN0 and mixed), definition of FDG-PET positive (quantitative, qualitative and unclear), study design (prospective or retrospective) and blinding (yes, no and not reported).

Results of the review
Thirty five studies were included in the review; 32 studies with a total of 1,236 participants were included in the meta-analyses. Sixteen studies were prospective, 10 retrospective and nine unclear. Six of the included studies were blinded, five were unblinded and the blinding status of 16 was unclear. Nineteen studies assessed the positivity of FDG-PET qualitatively and eight used quantitative definitions.

The overall sensitivity and specificity estimates for FDG-PET (32 studies) were 79% (95% CI 72 to 85) for sensitivity and 86% (95% CI 83 to 89) for specificity. Sensitivity was reduced for studies of cN0 patients only, 50% (95% CI 37 to 63; 10 studies, 311 participants) and for studies that reported blinded interpretation of FDG-PET, 72% (95% CI 51 to 87; five studies, 278 participants); specificity estimates were similar to that for the overall population in both cases. Study design and type of assessment of FDG-PET positivity did not significantly effect diagnostic accuracy.

When FDG-PET was compared with conventional diagnostic tests as a group (24 studies) and individually (16 studies for CT, nine studies for MRI, four studies for CT/MRI and four studies for USFNA), no significant differences were found. A comparison of FDG-PET with conventional diagnostic tests as a group in cN0 patients (five studies) found no significant differences.

Likelihood ratios for FDG-PET and conventional tests and NPVs for FDG-PET were also reported.

Authors’ conclusions
18F-FDG-PET had good diagnostic performance in the overall evaluation of patients with head and neck squamous cell carcinoma prior to surgical treatment. However, it failed to detect disease in half of patients with metastases and clinically negative cervical lymph nodes.

CRD commentary
The review addressed a clearly stated research question defined by appropriate inclusion criteria. The search strategy was limited to MEDLINE and reference screening, so some relevant studies may have been missed. No language restrictions were applied and a search strategy aimed at maximising sensitivity (based on index test and target condition only) was used. Measures to minimise error and bias were applied to the data extraction process, but it was unclear whether similar methods were used in the selection of studies for the review. The authors did not explicitly state that they assessed the methodological quality of included studies, but the presence of verification bias was used as an exclusion criterion and other aspects of study quality were extracted and their impact on test performance considered using subgroup analyses. Robust meta-analytic methods were used to synthesise data and despite some limitations in the review process the authors’ conclusions are likely to be reliable.

Implications of the review for practice and research
Practice: The authors stated that there was no solid evidence to support the routine clinical application of 18F-FDG-PET in the pretreatment evaluation of lymph node status in patients with HNSCC, including patients with clinically negative neck. They further stated that other imaging methods appeared to have similarly limited or worse diagnostic performance in these patients. Consideration of cost and potential complications versus potential information yield can be considered on a case-by-case basis.

Research: The authors stated that larger studies may clarify whether the use of 18F-FDG-PET offered an incremental improvement over conventional imaging methods.
Funding
Not reported.

Bibliographic details

PubMedID
18477804

DOI
10.1093/jnci/djn125

Original Paper URL
http://jnci.oxfordjournals.org/cgi/content/full/100/10/712

Indexing Status
Subject indexing assigned by NLM

MeSH
Confidence Intervals; Fluorodeoxyglucose F18; Head and Neck Neoplasms /pathology /radionuclide imaging; Humans; Lymph Nodes /pathology /radionuclide imaging; Lymphatic Metastasis; Neck; Neoplasm Staging; Positron-Emission Tomography /methods; Predictive Value of Tests; ROC Curve; Radiopharmaceuticals; Sensitivity and Specificity

AccessionNumber
12008103934

Date bibliographic record published
31/03/2009

Date abstract record published
16/12/2009

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.