Test performance of positron emission tomography and computed tomography for mediastinal staging in patients with non-small-cell lung cancer: a meta-analysis


CRD summary
This well-conducted review compared the accuracy of computed tomography (CT) and positron emission tomography (PET) with 18-fluorodeoxyglucose (FDG) for detecting the spread of non-small-cell lung cancer to mediastinal lymph nodes. FDG-PET was more accurate than CT, but because PET has more false-positive findings in people with enlarged nodes, positive findings need confirmation by biopsy. The authors' conclusions are reliable.

Authors' objectives
To compare the diagnostic accuracy of computed tomography (CT) and positron emission tomography (PET) with 18-fluorodeoxyglucose (FDG) for identifying mediastinal metastases in people with non-small-cell lung cancer (NSCLC). The authors also assessed whether the results of these tests are conditionally dependent, that is, whether the sensitivity and specificity of FDG-PET depends on whether enlarged mediastinal lymph nodes are identified on CT.

Searching
The authors searched Cancerlit (to June 2002), MEDLINE, EMBASE, Current Contents, BIOSIS Previews (to March 2003), and the reference lists of identified studies and reviews for studies published in any language. The search terms were available.

Study selection
Study designs of evaluations included in the review
Published full-length studies of any design that included at least 10 participants were eligible for inclusion.

Specific interventions included in the review
Studies that examined both FDG-PET imaging and CT were eligible for inclusion in the review; studies of CT alone were not included. Some of the included studies (10 out of 39) contained usable information only about FDG-PET.

Reference standard test against which the new test was compared
The authors did not report a specific reference test that studies had to include in order to be eligible for the review. More than 90% of the included studies required biopsy to confirm mediastinal metastasis, while 47% required thoracotomy with systematic sampling of both normal and abnormal appearing lymph nodes at all accessible mediastinal stations to verify the absence of lymph node involvement. In 56% of the included studies, those interpreting FDG-PET and CT findings were blinded to the final diagnosis.

Participants included in the review
Studies of people with known or suspected NSCLC were eligible if they included at least 10 participants, a minimum of five of whom had lymph node metastases. The mean age of the participants ranged from 56 to 69 years. The median proportion of females was 36% (range: 1 to 52). In studies that reported results using individuals as the unit of analysis, the median prevalence of malignant lymph nodes was 32% (range: 5 to 64).

Outcomes assessed in the review
Studies were eligible for inclusion if they provided sufficient data to permit the calculation of test sensitivity and specificity for identifying malignant lymph node involvement.

How were decisions on the relevance of primary studies made?
For studies reported in English, two authors independently evaluated papers for inclusion. Any disagreements were resolved by discussion. One author performed this task for studies in other languages. The authors were blinded to the journal, author, institutional affiliation and date of publication.
Assessment of study quality
The authors rated the methodological quality of the included studies using an adapted scale that focused on seven aspects: technical quality of the index tests, technical quality and application of the reference test, independence of test interpretation, description of study population, cohort assembly, sample size, and unit of analysis. For studies reported in English, two blinded investigators evaluated each study independently. For studies in other languages, one investigator performed the quality assessment.

Data extraction
Two investigators independently abstracted data about demographic characteristics, prevalence of malignant lymph node involvement, and the sensitivity and specificity of CT and FDG-PET for identifying malignant lymph nodes. In studies where individual patients were the unit of analysis, the authors extracted data about the ability of tests to distinguish ipsilateral or contralateral mediastinal lymph node involvement (N2 or N3) from hilar, intrapulmonary, or no lymph node involvement (N0 or N1). When it was not possible to make this distinction, the authors assessed test sensitivity and specificity for distinguishing N0 from N1, N2, or N3 lymph node status. In studies where individual patients were not the unit of analysis, the authors assessed sensitivity and specificity for identifying malignant lymph nodes or node stations. They analysed data from these studies separately.

The authors constructed a 2x2 contingency table for each study, whereby all participants were classified as having positive or negative results and as having or not having mediastinal lymph node involvement according to the reference test(s). True-positive and false-positive rates, log odds ratios for FDG-PET and CT, and 95% confidence intervals (CIs) were calculated.

Methods of synthesis
How were the studies combined?
The authors constructed summary receiver operating characteristic curves from summary diagnostic odds ratios and used a summary log odds ratio to describe trends. They used a fixed-effect model, or a random-effects model when there was evidence of statistical heterogeneity. The findings were expressed as a maximum joint sensitivity and specificity: the point on the summary receiver operating characteristic curve at which sensitivity and specificity are equal.

How were differences between studies investigated?
The authors appear to have conducted statistical tests for heterogeneity, although these were not reported in any detail. The sensitivity and specificity of the tests were compared according to the unit of analysis, year of publication, language and aspects of methodological quality using meta-regression. The sensitivity and specificity of the tests were also compared in participants with and without lymph node enlargement on CT, using discriminant function analysis and a non-parametric permutation test.

Results of the review
The review included 39 studies. The authors did not report the study types or overall participant numbers. However, 51% of the included studies were prospective.

Pooled data suggested that FDG-PET is more accurate than CT for identifying mediastinal lymph node involvement (P<0.001).

The median sensitivity of CT was 61% (interquartile range, IQR: 50, 71), while that of FDG-PET was 85% (IQR: 67, 91).

The median specificity of CT was 79% (IQR: 66, 89), while that of FDG-PET was 90% (IQR: 82, 96).

The maximum joint sensitivity and specificity of CT was 70% (95% CI: 67, 73), while that of FDG-PET was 86% (95% CI: 84, 88).
The results for CT were based on 1,119 participants from 23 studies, while those for FDG-PET were based on 1,959 participants in 32 studies.

Fourteen studies provided data about conditional sensitivity and specificity. FDG-PET was more sensitive, but less specific, when CT found enlarged lymph nodes (median sensitivity 100%, median specificity 78%) compared with when CT found no lymph node enlargement (median sensitivity 82%, median specificity 93%, \( P=0.002 \)).

The methodological quality of the studies varied, but the authors suggested that few aspects of methodological quality affected the reported diagnostic accuracy. Funnel plots found no evidence of publication bias.

**Authors' conclusions**
FDG-PET is more accurate than CT for mediastinal staging in people with potentially resectable NSCLC. The sensitivity and specificity of FDG-PET are related to the presence or absence of enlarged mediastinal lymph nodes on CT.

**CRD commentary**
This review had a defined research question and, generally, clearly described inclusion criteria. The exception was that no 'gold' standard reference test appears to have been specified as an inclusion criterion. The search strategy would appear appropriate and there were no language restrictions. Although the authors included only published full-length studies, funnel plots found no evidence of publication bias. The review was well reported, providing details of the study selection, quality assessment and data extraction processes, and using more than one investigator to make decisions. The review appear to have been well conducted overall, with effort made to reduce potential investigator bias.

The quantitative methods used to synthesise the data appear appropriate. The authors justified the use of all methods used and reported the statistical methods in some detail. Appropriately, they analysed data according to the unit of analysis in different studies. They also referred to suspected heterogeneity between the studies, but did not report a formal statistical test of heterogeneity. However, they did note that methodological features did not impact on the findings of individual studies.

The data presented support the authors' conclusions.

**Implications of the review for practice and research**
Practice: The authors stated that positive findings on FDG-PET should be confirmed by biopsy before excluding curative surgery as a treatment option. Negative FDG-PET findings should be interpreted in the light of an individual's pre-test probability of mediastinal metastasis and the presence of enlarged mediastinal nodes on CT.

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

**Funding**
US Veterans Affairs.

**Bibliographic details**

PubMedID
14644890

Original Paper URL
http://www.annals.org/cgi/content/full/139/11/879
Indexing Status
Subject indexing assigned by NLM

MeSH
Aged; Carcinoma, Non-Small-Cell Lung /diagnosis /pathology; Female; Fluorodeoxyglucose F18; Humans; Lung Neoplasms /diagnosis /pathology; Lymphatic Metastasis; Male; Mediastinal Neoplasms /secondary; Middle Aged; Neoplasm Staging /methods; ROC Curve; Radiopharmaceuticals; Sensitivity and Specificity; Tomography, Emission-Computed /standards; Tomography, X-Ray Computed /standards

AccessionNumber
12004008053

Date bibliographic record published
30/06/2005

Date abstract record published
30/06/2005

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.