Metastases from non-small cell lung cancer: mediastinal staging in the 1990s. Meta-analytic comparison of PET and CT

Dwamena B A, Sonnad S S, Angobaldo J O, Wahl R L

Authors' objectives
The authors compared the accuracy of positron emission tomography (PET) versus computed tomography (CT) for diagnosing mediastinal node metastases in people with non-small-cell lung cancer (NSCLC).

Searching
MEDLINE was searched for English language studies published between January 1990 and January 1998; the search terms were reported. The authors also searched the bibliographies of retrieved articles and recent journals for additional studies. There was no attempt to include unpublished studies.

Study selection
Study designs of evaluations included in the review
Original, peer-reviewed, full-length studies comparing CT and/or PET with an objective diagnostic standard were eligible for inclusion. Abstracts and early reports of data later published in full were excluded. The authors prioritised prospective studies, but did not exclude retrospective studies. All the studies included in the review were case series, 70% of which were prospective. The authors prioritised studies that included at least 35 people with disease and 35 people without disease, but did not exclude studies with smaller samples.

Specific interventions included in the review
The authors did not specify any inclusion criteria relating to the diagnostic reference standard. Some of the studies included in the review used surgery and long-term follow-up to assess the presence of breast cancer and survival rates. Studies of the diagnostic performance of CT and/or PET 2-[fluorine-18]fluoro-2-deoxy-D-glucose (FDG) PET were eligible for inclusion in the review if: an objective diagnostic standard was present; and established diagnostic criteria were used for abnormal test results (e.g. abnormal test lymph node of at least 10 mm in CT; abnormal lymph node uptake exceeding that of mediastinal blood pool in PET).

Studies of PET were prioritised if the type of scanner was third generation or later; the participants fasted for at least 4 hours before scanning; the dose of FDG was reported; and other technical details were reported in full.

Studies of CT were prioritised if they used a third- or fourth- generation scanner; had a section scanning time of 2 seconds or less; had a maximum section thickness and interval of 10 mm; the scanning area included from above the apices through the adrenal glands; and the criteria for interpretation were reported in full.

Reference standard test against which the new test was compared
Studies were eligible for inclusion if they established mediastinal nodal status with results from histologic samples obtained at mediastinoscopy, thoracotomy and/or autopsy. The authors prioritised studies where the index tests and reference standards were interpreted independently by masked investigators. To avoid verification bias, the reference test was applied regardless of the imaging results.

Participants included in the review
Studies of people with NSCLC were included. The mean age of the participants was 61 years (range: 57 to 67; the mean age was reported incorrectly as 21 years in the article). Where reported, the proportion of women ranged from 13 to 45%.

Outcomes assessed in the review
Studies were eligible for inclusion if they reported diagnostic accuracy in sufficient detail to allow the reconstruction of 2x2 contingency tables of the raw data (true-positive, true-negative, false-positive and false-negative results). The outcomes assessed in the review were diagnostic sensitivity, specificity and overall accuracy.
How were decisions on the relevance of primary studies made?
Two authors independently reviewed the articles to determine their eligibility for more detailed analyses. Any disagreements were resolved by repeat review and discussion.

**Assessment of study quality**
The authors developed criteria to assess validity prospectively, based on 5 standards for evaluating the quality of diagnostic studies. The criteria included: description and quality of the imaging procedure; technical quality of the reference test; uniform application of the reference test; independence of interpretation; clinical description and spectrum composition; prospective assembly of cohort; sample size; and adequate reporting of the results. The studies did not have to fulfil these criteria to be eligible for the review, nor were the findings weighted according to these criteria when pooling the results. Two authors independently assessed study quality. Any disagreements were resolved by discussion and consensus.

**Data extraction**
Two authors independently abstracted data on publication details, the number of participants, unit of analysis and outcomes. The authors were not blinded to the study findings or publication details when extracting the data. The authors recalculated sensitivity, specificity, accuracy and predictive values for each study, based on the reconstructed 2x2 contingency tables.

**Methods of synthesis**
How were the studies combined?
The sensitivity and specificity of the diagnostic tests were compared using summary receiver operating characteristic (ROC) curve analysis with logistic transformation, Student's t-test, and regression analyses.

How were differences between studies investigated?
Differences between the studies were described and the study characteristics were tabulated. The authors examined the slopes of regression lines fitted to summary ROC curves to assess differences between the studies and whether it was appropriate to use the log odds ratio as a summary statistic for comparing test performance.

**Results of the review**
Twenty-nine case series with internal controls (n=2,226) were included: 14 studies compared PET and CT in the same sample and 15 focused on CT alone.

The authors found that PET was more accurate than CT for diagnosing nodal metastases in NSCLC.

The mean sensitivity was 0.79 (range: 0.62 to 0.97; 95% confidence interval, CI: 0.76, 0.82) for PET and 0.60 (range: 0.25 to 0.89; 95% CI: 0.58, 0.62) for CT. The mean specificities were 0.91 (range: 0.79 to 0.99; 95% CI: 0.89, 0.93) and 0.77 (range: 0.44 to 0.95; 95% CI: 0.75, 0.79) for PET and CT, respectively.

The diagnostic accuracy of PET was 92%, while that of CT was 75%. The positive predictive values of PET and CT were 90% and 50%, respectively, while the corresponding negative predictive values were 93% and 85%. The log odds ratio was 3.77 (95% CI: 2.77, 4.77) for PET and 1.79 (95% CI: 1.49, 2.09, P for difference < 0.001) for CT.

A subgroup analysis by type of study, unit of analysis, sample size, publication year, geographical location and clinical context did not alter the findings.

**Authors' conclusions**
FDG PET is more accurate than CT for identifying mediastinal nodal metastases in people with NSCLC.
CRD commentary
The authors reported their research question clearly, along with general inclusion and exclusion criteria. However, the search strategy appears to have been somewhat narrow. The authors included only English language studies published in peer-reviewed journals. This may have been an attempt to ensure that the studies included in the review were of a high quality, but this strategy is likely to have limited the number of studies identified. Furthermore, narrowing the search strategy may not have actually improved the quality of the studies considered for inclusion. Only one database was searched, which further limits the scope of the review. The authors did not consider how language bias or publication bias may have affected the review's findings. These limitations mean that we cannot be confident that all the relevant studies were included in the review.

The authors provided detailed descriptions of the methods used to select and appraise the primary studies. These methods seemed reasonable. However, although the authors described criteria to assess the general quality and validity of the studies, these quality criteria were not exclusion criteria, merely signposts for the general quality of the studies included in the review. For example, the authors specified that high-quality studies should include at least 35 people with nodal metastases and 35 people without nodal metastases, yet some studies with much smaller samples were included in the review. The authors conducted subgroup analyses to compare the findings of studies with different sample sizes, but it may have been more appropriate to exclude smaller studies from the review.

The authors tabulated the characteristics of each study and used appropriate statistical techniques to pool the data. The subgroup analyses conducted appear to have been appropriate and the authors addressed their research question clearly. Overall, the authors' conclusions appear to be supported by the pooled data presented. However, as described already, this data may be limited in both quality and in scope, so the authors' conclusion may not be generalisable.

Implications of the review for practice and research
Practice: The authors suggested that FDG PET is significantly more accurate than CT for identifying mediastinal lymph node metastases in people with NSCLC.

Research: The authors suggested that future investigators should improve study design, especially regarding cohort assembly, blinded interpretation of the index and reference tests, and presentation of findings.

Bibliographic details

PubMedID
10551237

DOI
10.1148/radiology.213.2.r99nv46530

Original Paper URL
http://radiology.rsna.jnls.org

Indexing Status
Subject indexing assigned by NLM

MeSH
Carcinoma, Non-Small-Cell Lung /radiography /radionuclide imaging /secondary; Humans; Lung Neoplasms /pathology /radiography /radionuclide imaging; Lymphatic Metastasis; Mediastinum; Neoplasm Staging; ROC Curve; Sensitivity and Specificity; Tomography, Emission-Computed; Tomography, X-Ray Computed

AccessionNumber
11999009946
Date bibliographic record published
30/06/2004

Date abstract record published
30/06/2004

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.