Transthoracic needle aspiration biopsy for the diagnosis of localised pulmonary lesions: a meta-analysis

Lacasse Y, Wong E, Guyatt G H, Cook D J

Authors' objectives
To evaluate the accuracy of transthoracic needle aspiration biopsy (TNAB) for the diagnosis of solitary or multiple localised pulmonary lesions.

Searching
MEDLINE was searched on CD-ROM (from 1966 to 1996) for articles in the English language. The keywords used were 'lung neoplasma and needle biopsy', or 'coin lesion/diagnosis' (exploding all terms). The Index Medicus search, which was restricted to 1963 to 1965, used the headings 'biopsy', 'coin lesion, pulmonary' and 'lung neoplasmas'. The reference lists of relevant articles were checked for additional articles.

Study selection
Study designs of evaluations included in the review
Studies of consecutive series of patients, in which at least 90% of the patients received the reference standard test, were eligible for inclusion. Of the 48 included studies, 15 were prospective and 7 were blinded.

Specific interventions included in the review
Papers reporting on 'aspiration needles' (Chiba and Zavala types), 'small gauge biopsy needles' and 'cutting needles' were considered. Reports on aspiration biopsies performed with any type of radiological guidance were also included.

The following diagnostic categories were used: positive for malignancy; suspicious for malignancy; anything other than malignant; non-diagnostic; benign but non-specific; specific benign diagnosis (e.g. tuberculosis).

Reference standard test against which the new test was compared
Papers using resection specimen, biopsy of an adjacent site with tumour involvement, long-term follow-up, or culture as the reference standard were included. Concomitant clinical diagnosis and finding of typical malignant cells on TNAB were not considered an acceptable reference standard.

Participants included in the review
Studies of patients presenting with parenchymal pulmonary solitary or multiple localised lesions were included. Studies of patients with other malignancies were included if at least 90% of the patients had parenchymal lesions, or if subgroups were reported separately.

Outcomes assessed in the review
The included studies were required to report sufficient data to derive the operating characteristics of TNAB in terms of sensitivity (true-positive rate) and specificity (true-negative rate) for the diagnosis of malignancy, either by the number of patients or by the number of procedures.

How were decisions on the relevance of primary studies made?
Two reviewers successively applied the inclusion criteria to the titles and abstracts of all citations obtained. The reviewers were blinded to the authors' names, journal and year of publication. The full reports of eligible citations were retrieved and independently assessed by the same two reviewers. Any disagreement was resolved by consulting a third reviewer. Attempts were made to contact authors where there was insufficient information to decide on inclusion.

Assessment of study quality
Three important biases were considered:
The ‘work up’ bias, which is introduced by excluding patients from the analysis because they were not submitted to the reference standard procedure.

The ‘review’ bias, which is introduced when the test result was not verified by the reference standard procedure.

The ‘test review’ bias, which is introduced when the observers are aware of either the clinical condition of the patient or the final diagnosis.

The first two biases were considered as crucial and were therefore used as exclusion criteria. Two remaining validity criteria were applied: whether an independent blind comparison with the reference standard was performed, which limits the test review bias; and whether the evaluation was prospective or retrospective. The authors did not state how the papers were assessed for validity, or how many reviewers performed the validity assessment.

Data extraction
Two reviewers abstracted information from all of the papers selected for inclusion. Any disagreements were resolved by consensus. The agreement between the two primary reviewers was moderate (weighted kappa=0.42; 95% confidence interval, CI: 0.35, 0.50). The abstracted information included: the distribution of diagnoses; the size of pulmonary lesions; the location of lesions; the type of needle; the type of radiological assistance; whether a cytopathologist was consulted during the procedure; the reference standard used, including the number of patients who were given a final diagnosis by either method; the type and number of complications; the final result related to the accuracy of the test.

The authors of all the studies included were asked to supply missing information. A response rate of 85% was achieved.

Methods of synthesis
How were the studies combined?
Five meta-analyses were conducted using the following thresholds:

malignant versus all other categories;
malignant or suspicious versus all other categories;
suspicious versus all other categories (excluding the 'malignant' category);
benign versus all other categories; and
specific benign versus all other categories.

The sensitivity and specificity were calculated, weighted by the inverse of the estimate variance, and were pooled as described by Hasselblad and Hedges (see Other Publications of Related Interest) using a random-effects model. From the pooled sensitivity and specificity, the corresponding likelihood ratio (true-positive rate to false-positive rate, or sensitivity/(1-specificity)) was estimated. Subgroup analyses were conducted to identify the source of heterogeneity if the outcomes did not meet the criteria for homogeneity.

The incidence rates of complications, restricted to pneumothoraces requiring or not requiring chest drainage, were pooled similarly.

How were differences between studies investigated?
The random-effects model was used and a test for homogeneity was performed in order to disclose it and to investigate its potential sources. Receiver operating characteristic (ROC) curves (plots of the sensitivity versus specificity of individual studies) were plotted to examine visually whether heterogeneity existed among the study results. Both the sensitivity and specificity were tested for homogeneity using a chi-squared test.
Results of the review
Forty-eight studies, involving 9,047 patients or biopsies, were included.

From the pooled sensitivity and specificity corresponding to each diagnostic threshold, the associated positive likelihood ratios were derived. The likelihood ratio was 72 for malignant disease versus all other categories, 49 for malignant or suspicious versus all others, 15 for suspicious versus all categories but malignant, 0.07 for benign versus all others, and 0.005 for specific benign diagnosis versus all others. Differences in the methodological quality of the studies, needle types, or whether a cytopathologist participated in the procedure failed to explain the heterogeneity of the results. Given a 50% probability of malignancy prior to the TNAB, the post-test probabilities of malignancy upon receiving the results would be 99% for malignant, 94% for suspicious, 7% for non-specific benign, and 0.6% for benign with a specific diagnosis in 0.6%.

Authors' conclusions
Given the intermediate pre-test probabilities that would probably lead to performing a TNAB, the findings of 'malignant' or a specific diagnosis of a benign condition provided definitive results. Findings of 'suspicious' markedly increased the probability of malignancy, and 'benign' markedly decreased it, although it may not be considered definite.

CRD commentary
This was a well-performed review on the use of TNAB for the diagnosis of solitary or multiple localised pulmonary lesions. The review question was clear and the inclusion criteria seem to have been well chosen. The search strategy was limited, as only MEDLINE was searched and only English studies were included. In addition, no attempt to identify unpublished data was reported. It is possible that relevant data may have been omitted. The inclusion and exclusion of studies and the data extraction were performed in an acceptable manner, but it was not reported how the methodological quality assessment was conducted.

A summary estimate of effect across studies was generated. The method used was clearly described and justified, although a meta-analytic derivation of pooled likelihood ratio from individual study data may have been preferable to the calculation from pooled sensitivity and specificity; summary ROC curves could have been used to investigated the effects of differing diagnostic thresholds between the studies. Differences between the studies were assessed by a chi-squared analysis and by visual examination, and potential sources of heterogeneity were appropriately explored.

The conclusions of the review seem to follow from the evidence presented. The authors acknowledged that the critical appraisal of the studies included in the review revealed several limitations. For instance, only two studies were prospective and used a blind comparison with a reference standard. The limited inter-rater agreement for the inclusion of studies also reflected the often ambiguous and inconsistent reporting of the study methods.

Implications of the review for practice and research
Practice: The authors did not state any implications for practice.

Research: The authors stated that the heterogeneity of the results suggests that local studies would be highly desirable to appreciate more the operating characteristics of the tests within institutions. Inter-observer agreement, technical aspects of the biopsy, clinical indications, and cost analysis also need to be addressed.

Bibliographic details

PubMedID
10491450
Original Paper URL
http://thorax.bmjournals.com/cgi/content/full/54/10/884

Other publications of related interest

Indexing Status
Subject indexing assigned by NLM

MeSH
Biopsy /adverse effects /methods; Humans; Lung /pathology; Lung Neoplasms /pathology; Sensitivity and Specificity; Solitary Pulmonary Nodule /pathology

AccessionNumber
11999001940

Date bibliographic record published
30/04/2004

Date abstract record published
30/04/2004

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.