

LUNG CANCER

Accuracy of transbronchial needle aspiration for mediastinal staging of non-small cell lung cancer: a meta-analysis

J-E C Holty, W G Kuschner, M K Gould



Thorax 2005;60:949-955. doi: 10.1136/thx.2005.041525

See end of article for authors' affiliations

Correspondence to: Dr J-E C Holty, Center for Primary Care and Outcomes Research, Stanford University, 117 Encina Commons, Stanford, CA 94305-6019, USA; jholty@stanford.edu

Received 27 January 2005
Accepted 8 June 2005
Published Online First
30 June 2005

Background: The reported accuracy of transbronchial needle aspiration (TBNA) for mediastinal staging in non-small cell lung cancer (NSCLC) varies widely. We performed a meta-analysis to estimate the accuracy of TBNA for mediastinal staging in NSCLC.

Methods: Medline, Embase, and the bibliographies of retrieved articles were searched for studies evaluating TBNA accuracy with no language restriction. Meta-analytical methods were used to construct summary receiver-operating characteristic curves and to pool sensitivity and specificity.

Results: Thirteen studies met inclusion criteria, including six studies that surgically confirmed all TBNA results and enrolled at least 10 patients with and without mediastinal metastasis (tier 1). Methodological quality varied but did not affect diagnostic accuracy. In tier 1 studies the median prevalence of mediastinal metastasis was 34%. Using a random effects model, the pooled sensitivity and specificity were 39% (95% CI 17 to 61) and 99% (95% CI 96 to 100), respectively. Compared with tier 1 studies, the median prevalence of mediastinal metastasis (81%; $p=0.002$) and pooled sensitivity (78%; 95% CI 71 to 84; $p=0.009$) were higher in non-tier 1 studies. Sensitivity analysis confirmed that the sensitivity of TBNA depends critically on the prevalence of mediastinal metastasis. The pooled major complication rate was 0.3% (95% CI 0.01 to 4).

Conclusions: When properly performed, TBNA is highly specific for identifying mediastinal metastasis in patients with NSCLC, but sensitivity depends critically on the study methods and patient population. In populations with a lower prevalence of mediastinal metastasis, the sensitivity of TBNA is much lower than reported in recent lung cancer guidelines.

Non-small cell lung cancer (NSCLC) is the most common malignancy in the world and accounts for an estimated 1 million deaths each year.¹ The overall 5 year survival is approximately 15%.² However, the survival rate approaches 70% in some patients with resectable disease.³ Metastasis to the mediastinal lymph nodes is one of the most important factors in determining resectability and prognosis.⁴ Careful mediastinal staging is essential to identify appropriate candidates for surgery and to avoid futile thoracotomy in patients with more advanced disease.

Currently, computed tomography (CT) is the most frequently used preoperative staging modality. However, large benign hyperplastic lymph nodes are commonly found in patients with NSCLC⁵ and normal sized lymph nodes frequently harbour metastases.⁶ Preoperative clinical staging with CT differs from surgical staging in as many as 40% of cases.⁷⁻⁸ Newer imaging modalities such as positron emission tomography (PET) have limitations in diagnostic accuracy as well.⁹ Given the limitations of CT and PET, invasive surgical staging techniques such as mediastinoscopy are often used to exclude or confirm mediastinal lymph node metastasis, especially in patients who are candidates for surgical resection. However, mediastinoscopy is associated with a complication rate of 2-3% and a surgical mortality rate of around 0.1%.¹⁰⁻¹²

Transbronchial needle aspiration (TBNA) using a flexible fiberoptic bronchoscope was developed in the early 1980s to obviate the need for more risky surgical staging procedures. Compared with mediastinoscopy, TBNA is generally more convenient, less risky, and less expensive.¹³ A recent systematic review of mediastinal staging with TBNA found the sensitivity to be similar to that of mediastinoscopy (76% v

81%).¹⁴ This analysis, however, was not restricted to patients with NSCLC, did not assess study methodological quality, and did not attempt to identify sources of variation in study results. We performed a meta-analysis to estimate the diagnostic accuracy of TBNA in patients with NSCLC and to identify technical factors and patient characteristics that have an impact on accuracy.

METHODS

A more detailed description of our methods is available as an online data supplement on the *Thorax* website at www.thoraxjnl.com/supplemental.

Literature search and identification of studies

Medline and Embase (January 1966 to July 2003; Medline updated through April 2004) were searched to identify studies that examined TBNA for mediastinal staging in NSCLC (fig S1A and B, online supplement), and reference lists of included studies and review articles were manually searched. All articles were considered, regardless of language.

Selection of studies

We included studies that (1) examined TBNA using a flexible bronchoscope for mediastinal staging in patients with NSCLC; (2) enrolled at least 10 subjects with and/or 10 subjects without mediastinal metastasis; (3) provided sufficient data to permit calculation of sensitivity and/or specificity; and (4) enrolled no more than 10% of patients with a diagnosis other than NSCLC or provided separate data

Abbreviations: NSCLC, non-small cell lung cancer; TBNA, transbronchial needle aspiration

for patients with NSCLC. More rigorous (tier 1) studies enrolled at least 10 subjects with and 10 subjects without mediastinal lymph node involvement, surgically confirmed all TBNA results (for example, with mediastinoscopy, mediastomy and/or thoracotomy), and used the patient as the unit of analysis. The authors of abstracts and studies not reporting sufficient data were contacted to request additional information.

Study quality

An existing instrument was adapted to describe the methodological quality of studies,¹⁵ as reported previously (fig S2, online supplement).⁹ ¹⁶ We developed criteria for the technical quality of TBNA based on our clinical experience and by reviewing published guidelines.^{17–19}

Data abstraction

One investigator abstracted primary data regarding patient characteristics and the sensitivity and/or specificity of TBNA for identifying mediastinal metastasis in patients with NSCLC.

When possible, we separated staging characteristics of TBNA for patients with and without enlarged lymph nodes on the CT scan and for biopsies performed at hilar, subcarinal, paratracheal, or other lymph node stations. We also separately tabulated test characteristics for studies using “real time” imaging—for example, CT fluoroscopy, endobronchial ultrasound, or transthoracic ultrasound.

Data synthesis and statistical analysis

We constructed a 2×2 contingency table for each study to summarise the results of TBNA and the reference test(s). For each study we calculated the true positive rate (TPR; sensitivity), the false positive rate (FPR; 1–specificity), and the log odds ratio (LOR). When necessary, we added 0.5 as a correction factor to calculate the LOR.

Because many studies of TBNA did not confirm positive test results surgically, they were unable to report false positive rates. We therefore calculated a weighted kappa-1 coefficient which does not require information about the false positive rate to measure test accuracy with respect to avoiding false negative results.^{20–21}

A random effects model was used to pool sensitivity, specificity, LOR and kappa-1.²² When pooling sensitivity and specificity, we excluded studies with <10 subjects with or without mediastinal lymph node involvement, respectively, in the calculations. Summary receiver operating characteristic (SROC) curves as described by Moses *et al*²³ were constructed to summarise the results quantitatively.

To assess sources of variation in study results we performed sensitivity analyses, discriminate function analyses, and meta-regressions. Sensitivity analysis included stepwise single study elimination, adjusting the correction factor, and varying the reference test result in studies that employed a suboptimal reference standard. To compare sensitivity and specificity jointly in studies grouped by tier and prevalence we used discriminant function analysis. Multivariate analysis of variance (ANOVA)^{24–25} was used to compare reported sensitivities and LORs in studies with high and low prevalences of lymph node metastasis ($\geq 60\%$ or $< 60\%$) and year of study publication (≥ 1995 or < 1995). To assess for the presence of publication bias we constructed inverted funnel plots of standard error versus estimated effect size (LOR) for each individual study.²⁶ We also assessed how the exclusion of small cell cancer cases from the included studies impacted on the accuracy of TBNA.

All biostatistical models were programmed with Excel 8.0 for Windows (Microsoft Corporation, Redmond, Washington, USA). Discriminant function analysis was performed using SAS 9.0 for Windows (SAS Corp, Cary, NC, USA). We calculated 95% confidence intervals (CIs) for the TPR and the FPR by using the quadratic method.²⁷ A normal approximation to the binomial of the standard error was used in calculating all other confidence intervals, as appropriate. When making comparisons between groups of studies, an unpaired *t* test or the Mann-Whitney U test was used, as appropriate. A two tailed *p* value of < 0.05 was considered statistically significant.

RESULTS

Literature search and study selection

Our literature search identified 525 potentially eligible studies (fig 1); 398 studies judged not to be relevant after carefully reviewing their titles and abstracts were eliminated. A hand search of the bibliographies of the remaining 127 articles identified 203 additional studies that were potentially relevant. A preliminary review of these 330 articles eliminated 268 studies, leaving 62 articles for detailed analysis (table S1, online supplement). After detailed review, 13 studies met the inclusion criteria (table 1).^{28–40} Studies were most often excluded because they provided insufficient data to calculate sensitivity or specificity (76%) or enrolled more than 10% of subjects with a diagnosis other than NSCLC (60%). Inter-rater agreement for study inclusion was high (mean kappa $\sim 80\%$; table S2, online supplement). Five authors provided additional information that enabled us to include their studies.^{28–29–32–33–37}

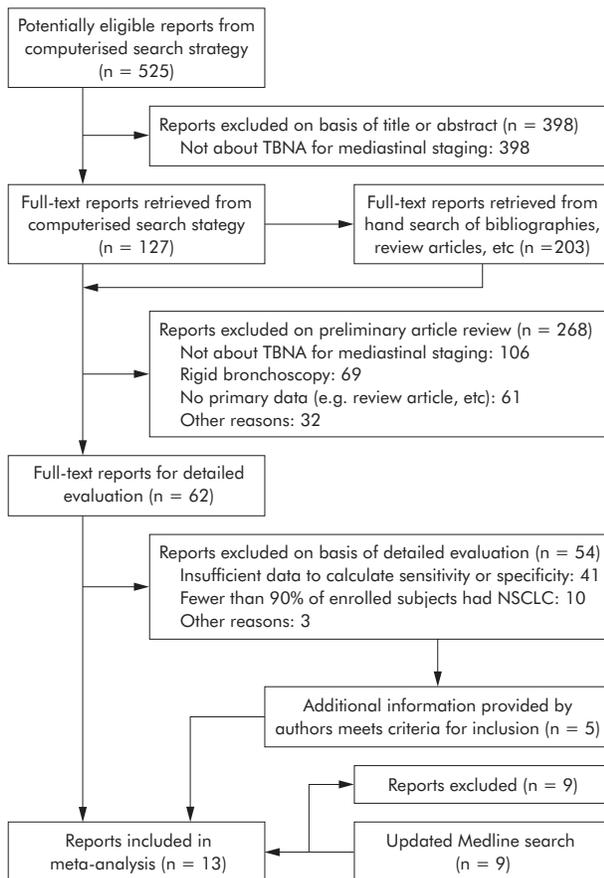


Figure 1 Literature search and selection. Studies could meet one or more exclusion criteria. For simplicity, only one primary exclusion criterion per study is shown.

Table 1 Characteristics of studies in the meta-analysis*

Study (reference)	Year	Prevalence [§]	NSCLC (%)	TBNA results ^{‡‡} (%)				Inclusion criteria and comments
				TP	FN	FP	TN	
<i>Tier 1 studies</i> [†] Harrow <i>et al</i> ^{†‡}	2000	34	100	8	17	1	48	Patients with suspected lung cancer were included. Patients without lung cancer, or TBNA from a lymph node confluent with the tumour mass were excluded by the authors. Patients with potentially resectable extrabronchial or endobronchial mass suggestive of lung cancer and without extrathoracic metastases were included. We excluded patients with N0 or N1 disease on pre-TBNA CT.* Patients with potentially resectable lung cancer without evidence of extrathoracic metastases and with enlarged mediastinal LAD on CT or CXR were included. Patients with potentially resectable lung cancer (no preoperative SCLC) without evidence of extrathoracic metastases were included. Only subcarinal lymph nodes were biopsied. Patients with CXR evidence of lung cancer and who were potentially resectable without evidence of extrathoracic metastases were included. One FP had scanty neoplastic cells without lymphocytes.
Bilaceroglu <i>et al</i> ^{†‡}	1998	60	100	24	9	0	22	
Disdier <i>et al</i> ^{†‡}	1998	52	100	5	9	1	12	
Ratto <i>et al</i> ^{†‡}	1988	30	96	2	12	0	33	
Schenk <i>et al</i> ^{†‡}	1986	29	100	8	13	2	50	
<i>Studies not meeting tier 1 criteria</i> Herth <i>et al</i> ^{†‡} **	2002	82	100	30	6	0	8	Patients with central lung cancer and enlarged mediastinal LAD by CT without evidence of extrathoracic metastases were included. Needle placement via endobronchial ultrasound. Inclusion criteria not specified. Unable to calculate specificity (100% prevalence). Patients with NSCLC, N2 disease by CT scanning and a negative bronchoscopy for lung cancer were included. Statistical analysis was by TBNA specimen. Data were collected by retrospective chart review. Non-surgically confirmed negative TBNA specimens (total 49) were assumed to be false negative. Patients with CXR evidence of lung cancer and enlarged mediastinal LAD on CT and who were potentially resectable and without evidence of extrathoracic metastases were included. Patients undergoing thoracotomy for mediastinal adenopathy on CT and suspected lung cancer were included. Bronchoscopist was not experienced with TBNA. Real-time CT assisted needle placement. Patients with resectable lung cancer and paratracheal LAD without extrathoracic metastases were included. Four of the 32 TPs were confirmed surgically. Patients with CXR evidence of lung cancer and who were potentially resectable and without evidence of extrathoracic metastases were included. Patients with suspected lung cancer and who were potentially resectable and without evidence of extrathoracic metastases were included.
Wang <i>et al</i> ^{†‡}	2002	100	93 ^{††}	10	0	0	0	
Patelli <i>et al</i> ^{†‡}	2002	NA	100	127	52	–	15	
Katis <i>et al</i> ^{†‡}	1998	95	100	28	8	–	2	
Rong <i>et al</i> ^{†‡} **	1998	79	100	26	2	0	5	
Schenk <i>et al</i> ^{†‡}	1993	81	100	32	6	–	9	
Schenk <i>et al</i> ^{†‡}	1989	81	100	14	3	–	4	
Wang <i>et al</i> ^{†‡}	1983	55	100	13	3	–	13	

TBNA, transbronchial needle aspiration; NSCLC, non-small cell lung cancer; SCLC, small cell lung cancer; LAD, lymphadenopathy; CT, chest computed tomography; CXR: chest radiograph; NA, not applicable; TP, true positive; FN, false negative; FP, false positive; TN, true negative; –, studies in which positive TBNA results were not confirmed surgically but were assumed to be true positives.

*Five studies^{30 31 33 34 36} reported age (median 60 years) and six studies^{30 31 33–36} reported sex characteristics (median proportion male 91%).

[†]Studies meeting tier 1 criteria surgically confirmed all TBNA had at least 10 subjects with and without mediastinal lymph node metastasis and used the patient as the unit of analysis.

[‡]Additional information obtained from original study author(s).

[§]Prevalence of mediastinal lymph node metastasis. Studies that did not surgically confirm all TBNA results assumed that the false positive rate was zero.

*All mediastinal and hilar lymph nodes (≥8 mm) identified by CT scanning in this study were sampled by TBNA. To estimate TBNA diagnostic accuracy for identifying mediastinal metastasis, patients with N0 or N1 disease on pre-TBNA CT scanning were excluded from the analysis (the combined mediastinal/hilar TBNA results for NSCLC were: TP 49, FN 18, FP 0, TN 16). The sensitivity (73% v 73%; p=0.97) and specificity (100% v 100%; p=1.0) were similar whether or not N0/N1 disease on CT scanning was excluded.

**Two studies used real-time radiological assistance to guide needle placement during TBNA.

^{††}Three of 42 patients undergoing TBNA had SCLC. Thirty two patients who had hilar lymph node biopsies (not mediastinal) were excluded.

^{‡‡}Number of patients (except for the study by Patelli *et al*,³⁵ where statistical analysis was by TBNA specimen).

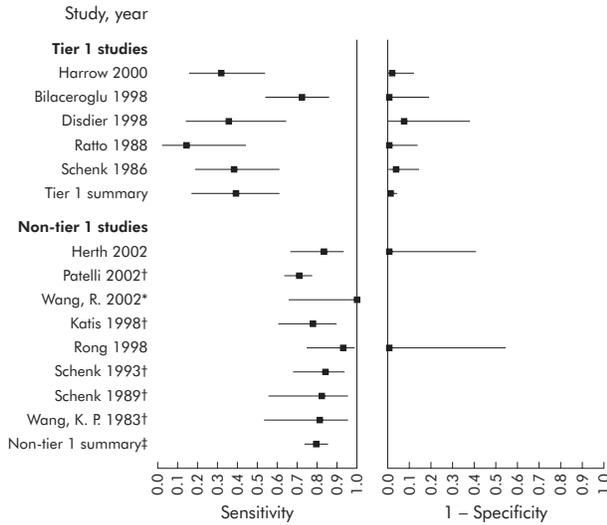


Figure 2 Individual study estimates of sensitivity and 1 – specificity of TBNA for identifying mediastinal metastasis. Error bars represent 95% confidence intervals (CIs). Point estimates and 95% CIs for tier 1 studies and studies meeting inclusion criteria but not tier 1 criteria are shown. Tier 1 and non-tier 1 summary point estimates and corresponding 95% CIs are shown and were calculated using a random effects model. *Specificity was not calculated for the study by Wang R *et al*⁴ because the prevalence of mediastinal lymph node metastasis was 100%. †Specificity is not shown for studies that did not surgically confirm all TBNA results but instead assumed that all positive results were true positives. ‡Summary non-tier 1 – specificity not shown because only two studies allowed calculation of specificity (both 100%),^{33,37} although both of these studies had fewer than 10 patients without mediastinal metastases.

Study description

The median number of participants per study was 44 (range 10–183). Six studies^{30,31,33,34,36,37} reported statistics about the age of participants (median age 60 years) and seven studies^{30,31,33–37} reported sex characteristics (median proportion male 89%). One study reported results by using individual lymph nodes as the unit of analysis.³⁵ For the other studies that reported results by using the patient as the unit of analysis, the median prevalence of mediastinal metastasis was 70% (interquartile range 47–83). The size and type of TBNA needle used and the number of aspirate passes per lymph node station varied between studies (table S3, online supplement). None of the studies stratified results according to nodal station or lymph node size on the CT scan in patients with NSCLC. In eight studies all positive and negative TBNA results were confirmed by mediastinoscopy, mediastinotomy, or thoracotomy.^{28–34,37} Six studies enrolled fewer than 10 subjects without mediastinal lymph node involvement.^{33,34,36–39} Two studies^{33,37} used real-time imaging (CT or endobronchial ultrasound) to guide needle placement during TBNA. Five studies met criteria for tier 1 analysis.^{28–32}

Study quality

Studies met between 12 and 23 of the 34 prespecified criteria for methodological quality. Seven studies met at least 50% of the criteria.^{28–30,32,33,36,38} Table S4 (online supplement) shows selected aspects of methodological quality for each study. In general, tier 1 studies met more criteria (mean 18.8; 95% CI 15.8 to 21.8) than non-tier 1 studies (mean 15.8; 95% CI 14.0 to 17.5), but this difference was not statistically significant ($p = 0.13$).

Diagnostic accuracy of TBNA

Tier 1 analysis (5 studies)

In these studies the median prevalence of mediastinal metastasis was 34% (range 29–60). The median sensitivity

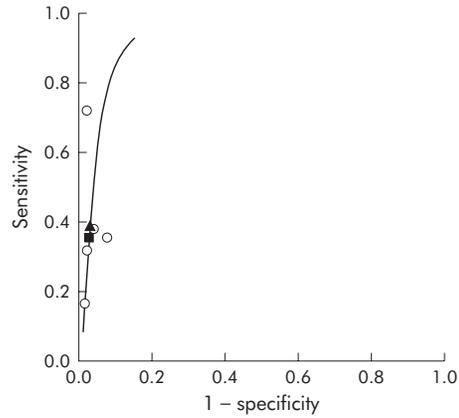


Figure 3 Summary receiver-operating characteristic curve for mediastinal staging with TBNA for tier 1 studies. Individual study estimates of sensitivity and 1 – specificity are shown (open circles). Median (solid square) and pooled sensitivities (solid triangle) are also shown.

and specificity of TBNA were 36% (interquartile range 32–38) and 98% (interquartile range 96–100), respectively (table 2). The pooled (random effects) sensitivity was 39% (95% CI 17 to 61) and the pooled specificity was 99% (95% CI 96 to 100) (table 2, fig 2). The corresponding positive and negative likelihood ratios were 29.0 and 0.62, respectively. The summary ROC curve is shown in fig 3.

The pooled (random effects) kappa-1 coefficient was 30% (95% CI 15 to 46), suggesting that the accuracy of TBNA with respect to false negative results was poor to fair in tier 1 studies.

Non-tier 1 analysis (8 studies)

Two non-tier 1 studies used real-time radiological needle guidance during TBNA.^{33,37} In the remaining six studies the median prevalence of mediastinal metastasis was 81% (range 55–100; $p = 0.002$ for comparison with tier 1 studies). None of these six studies provided sufficient information to calculate specificity (for example, they did not surgically confirm positive TBNA results). The median sensitivity of TBNA in studies not using real-time radiological needle guidance was 82% (interquartile range 79–84; table 2). The pooled (random effects) sensitivity was 78% (95% CI 71 to 84; table 2, fig 2). The pooled kappa-1 coefficient (random effects) was 40% (95% CI 19 to 62; table 2), suggesting that the accuracy of TBNA with respect to false negative results was fair in non-tier 1 studies.

The median prevalence of mediastinal metastasis in the two non-tier 1 studies that used real-time radiological needle guidance was 83% ($p = 0.84$ for comparison with the six other non-tier 1 studies). The pooled (85%) and median sensitivities (88%) in these two studies were not significantly different ($p = 0.36$ and $p = 0.38$, respectively) from the pooled and median sensitivities of the six non-tier 1 studies that did not use real-time radiological guidance.

Summary analysis (11 studies)

The Q statistic from the random effects model showed that there was statistically significant heterogeneity in sensitivity ($p < 0.001$) but not in specificity ($p = 0.90$). Discriminant function analysis confirmed that there was a statistically significant difference in the joint sensitivity and specificity of tier 1 and non-tier 1 studies ($p = 0.002$, parametric Wilks’ lambda test; fig 4). We therefore did not pool the results of tier 1 and non-tier 1 studies.

Table 2 Summary of results*

	Median sensitivity (IQR)	Median specificity (IQR)¶	Pooled sensitivity (95% CI)†	Pooled specificity (95% CI)† ¶	Likelihood ratio (95% CI)‡ ¶	
					Positive	Negative
Tier 1 studies	0.36 (0.32–0.38)	0.98 (0.96–1.00)	0.39 (0.17 to 0.61)	0.99 (0.96 to 1.00)	29.0 (21.9 to 38.4)	0.62 (0.53 to 0.72)
Non-tier 1 studies	0.82 (0.79–0.84)	–	0.78 (0.71 to 0.84)	–	–	–
p values§	0.001	–	0.009	–	–	–

IQR, interquartile range.

*For studies that did not use real-time radiological guidance for needle placement.

†Pooled sensitivity and specificity were calculated using a random effects model.

‡Positive and negative likelihood ratios were calculated from the pooled sensitivity and specificity.

§p value for the comparison between tier 1 and non-tier 1 studies.

¶Median and pooled specificities for non-tier 1 studies were not calculated because only two studies allowed calculation of specificity (both 100%);^{33, 37} however, both of these studies had fewer than 10 patients without mediastinal metastases. We were thus unable to calculate likelihood ratios for non-tier 1 studies.

Complication rate

One study did not report complications.³⁰ Of the remaining studies, two reported major complications,^{28, 40} including two major bleeds and one pneumothorax requiring a chest tube. Two other cases of pneumothoraces³⁵ and one case of pneumomediastinum²⁸ spontaneously resolved under observation. The mean rate of major complications per patient in tier 1 and non-tier 1 studies was 0.32% (95% CI 0.01 to 6) and 0.25% (95% CI 0.01 to 6), respectively (p = 0.65). The overall major complication rate was 0.26% (95% CI 0.01 to 4).

Sensitivity analysis and meta-regressions

An inverted funnel plot showed no evidence of publication bias (fig S4, online supplement). Stepwise single study elimination did not substantially affect the magnitude of the pooled LOR or sensitivity in tier 1 or non-tier 1 studies (table S5, online supplement). In one study, one of two false positive results had scanty neoplastic cells and no lymphocytes.³² Re-categorising this result as a true negative had no effect on pooled sensitivity, specificity, LOR, or the kappa-1 coefficient. Varying the correction factor from 0.5 to 0.1 had no impact on the LOR or the kappa-1 coefficient. Using a 0.1 correction tended to shift the summary ROC curve to the left (increasing specificity), but had little discernable impact on sensitivity.

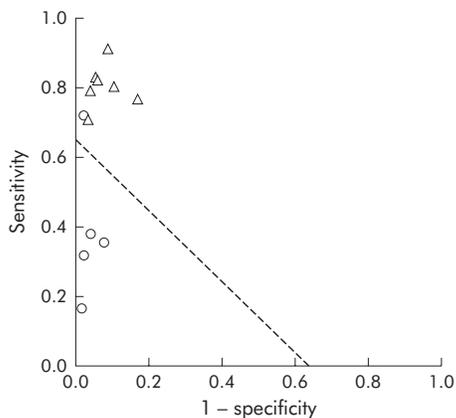


Figure 4 Discriminant function analysis for mediastinal staging with TBNA. Individual study estimates of sensitivity and 1 – specificity are shown for tier 1 studies (circles) and all other studies meeting inclusion criteria (triangles). The one study with 100% prevalence of mediastinal lymph node metastasis was not included (that is, specificity undefined).³⁴ The discriminant function that separates tier 1 from all other included studies is shown by the broken line. This confirms that differences in diagnostic accuracy between tier 1 and non-tier 1 studies are statistically significant when sensitivity and specificity are considered jointly (p=0.002, parametric Wilks’ lambda test). The one tier 1 study²⁹ that fell on the “wrong” side of the line had a relatively high prevalence (60%) of mediastinal metastasis.

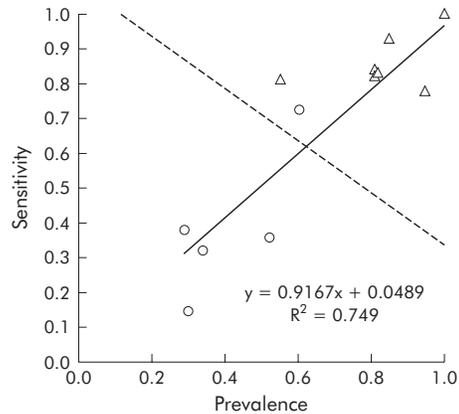


Figure 5 Sensitivity as a function of the prevalence of mediastinal lymph node metastasis. Individual study estimates of prevalence and sensitivity are shown for tier 1 (circles) and non-tier 1 (triangles) studies. Both the linear regression equation and R² are shown. One study reported results by using individual lymph nodes as the unit of analysis and was not included (unable to calculate prevalence of mediastinal metastasis).³⁵ The discriminant function that separates tier 1 from non-tier 1 studies (broken line) was significant (p=0.002, parametric Wilks’ lambda test).

Study sensitivity was positively correlated with the prevalence of lymph node metastasis (fig 5). When the prevalence rose from 40% to 80%, sensitivity increased from 42% to 78%. For the seven studies in which the prevalence of mediastinal disease was ≥60%, the median sensitivity (83% v 36%; p = 0.005) and pooled sensitivity (84% v 40%; p = 0.005) were higher than the five remaining studies in which prevalence was <60% (fig S3, online supplement). Discriminant function analysis confirmed that the joint sensitivity and specificity were different in studies with high versus low prevalence (p = 0.01, parametric Wilks’ lambda test).

For the eight studies published since 1995, the pooled sensitivity (71% v 60%; p=0.52) was not significantly different from the five remaining studies published before 1995. However, the median prevalence of lymph node metastasis in more recent studies (82% v 55%; p=0.09) was higher than in the five earlier studies.

These and other potential sources of heterogeneity were assessed by a multivariate ANOVA to compare reported sensitivities and LORs in studies with respect to the prevalence of lymph node metastasis (≥60% or <60%) and year of publication (≥1995 or <1995). Because only two included studies used real-time radiological needle guidance, we were unable to assess this potential source of heterogeneity and excluded these two studies from the analysis. Sensitivity was higher in studies with a higher prevalence of

lymph node metastasis (difference 60%; 95% CI 51 to 69) and in more recently published studies (difference 10%; 95% CI 1 to 18). The prevalence of lymph node metastasis, but not year of publication, had a significant effect on the LOR.

Excluding patients with small cell lung cancer from the included studies had no impact on the pooled sensitivity in tier 1 (39% *v* 41%, $p = 0.92$) or non-tier 1 (78% *v* 80%, $p = 0.71$) studies.

DISCUSSION

TBNA is highly specific for identifying mediastinal metastasis in patients with NSCLC, but sensitivity depends critically on the prevalence of mediastinal disease. Specificity is excellent, but not perfect. In three of eight studies that surgically confirmed all TBNA results, four false positive results were reported. One of the four false positive results would have been avoided if biopsy specimens were considered negative when they lacked nodal tissue or when the cytopathologist identified the specimen as “contaminated” or containing “atypical” cells. It is essential to avoid contamination of the bronchoscope channel and to follow stringent criteria to define positive or negative biopsy specimens in order to minimise the risk of false positive TBNA results. We found that TBNA is generally safe with a major complication rate of approximately 0.3%.

We identified several sources of variation in study results. Sensitivity was much lower in tier 1 studies than non-tier 1 studies. Tier 1 studies surgically confirmed all TBNA results, enrolled at least 10 patients with and without mediastinal metastasis, and used the patient as the unit of analysis. Sensitivity was also lower in studies with a low prevalence (<60%) of mediastinal metastasis. Not surprisingly, TBNA appears to be less sensitive than mediastinoscopy for identifying mediastinal metastasis. A recent meta-analysis of 14 studies of mediastinoscopy reported a pooled sensitivity of 81% (95% CI 76 to 85).¹⁴ In these studies the pooled prevalence of mediastinal disease was 37%, which is similar to the median prevalence (34%) of lymph node metastasis in tier 1 studies of TBNA.

The difference in diagnostic accuracy between tier 1 and non-tier 1 studies was statistically significant. We believe that this difference is probably related to a lower prevalence of mediastinal metastasis in tier 1 than in non-tier 1 studies. Higher disease prevalence and enrolment of patients with a more severe spectrum of disease are sources of variation in studies of diagnostic accuracy leading to an increase in sensitivity.^{41, 42} We speculate that the higher prevalence of mediastinal metastasis in non-tier 1 studies may reflect enrolment of study cohorts with a more severe spectrum of mediastinal disease, resulting in more positive TBNA results. For example, non-tier 1 (high prevalence) studies may have enrolled a greater number of patients with bulky lymphadenopathy in whom TBNA was being used to confirm the diagnosis of unresectable disease. In contrast, tier 1 (lower prevalence) studies may have enrolled potential surgical candidates with less impressive lymph node enlargement. A recent meta-analysis of 39 studies comparing PET with CT scanning for mediastinal staging in NSCLC found that the median prevalence of malignant lymph nodes in enrolled patients was 32% (range 5–64), which is similar to the median prevalence of mediastinal metastasis in the tier 1 studies in our analysis.⁹ Most of the studies of PET and CT scanning enrolled patients with potentially resectable NSCLC. Furthermore, the bronchoscopist’s technique may be more proficient when the pretest probability of obtaining a positive result is high (higher prevalence of mediastinal disease within the study cohort). For example, more diligence may be taken to identify endobronchial landmarks, more TBNA needle passes attempted, and more aggressive sedation given

to minimise cough and patient movement during the procedure.

The difference in pooled sensitivities between tier 1 and non-tier 1 studies may also be the result of methodological differences. Non-tier 1 studies used suboptimal methodological criteria by not confirming all TBNA results against a reference standard (verification bias), having insufficient numbers of participants with and without mediastinal metastasis, and/or not using the patient as the unit of analysis. Verification bias has been shown to lead to overestimates of test sensitivity.⁴¹

A previous meta-analysis showed that the pooled sensitivity of 12 studies analysing TBNA in patients with either small-cell lung cancer or NSCLC was 76%.¹⁴ Our estimates of sensitivity were lower for tier 1 studies (39%) because several studies that were included in this previous meta-analysis did not meet the criteria for our tier 1 analysis. Interestingly, the exclusion of patients with small cell lung cancer from the studies included in our analysis did not significantly affect sensitivity.

Despite the relatively low sensitivity of TBNA in detecting mediastinal metastasis compared with other invasive staging procedures, TBNA continues to be an appropriate diagnostic test in the sampling of mediastinal lymph nodes, especially if concurrently performed with routine bronchoscopic examination for suspected lung cancer. TBNA is generally more convenient, less risky, and less expensive than other invasive staging procedures such as mediastinoscopy.¹³ A formal assessment of the cost effectiveness of staging TBNA is beyond the scope of this analysis.

Although we were unable directly to assess how newer needles, use of on-site cytological analysis, and/or improved techniques may impact on TBNA accuracy, our multivariate ANOVA showed that more recent studies—which presumably used more up to date techniques and equipment—had a slightly higher sensitivity when we controlled for prevalence of mediastinal metastasis.

Our study has several limitations. Firstly, only a small number of studies met our inclusion criteria (five tier 1 and eight non-tier 1 studies). Most studies enrolled fewer than 100 participants and were performed at single centres where experience with TBNA is likely to be extensive. Large multicentre prospective studies of TBNA should be performed in consecutively enrolled patients with NSCLC. Studies should explicitly define inclusion criteria and should report separate results for patients with non-bulky and bulky lymphadenopathy. Secondly, because needle type and size, as well as the number of aspiration passes varied between studies, we were unable to control for these test characteristics. Likewise, because most studies did not report age or sex characteristics, we were unable to control for these demographic features. Thirdly, few of the included studies provided information on whether TBNA results altered patient management. Clearly, positive results on TBNA obviate the need for mediastinoscopy because specificity and positive predictive value are high. However, simple calculations based on our results indicate that, when prevalence is relatively low (~35%), approximately 85% of patients will have negative TBNA results and 25% of such results will be falsely negative. Fourthly, despite an exhaustive search, we may not have identified all studies, especially those with unpublished results. We identified one potentially relevant abstract but we were unable to obtain sufficient additional information to assess it for inclusion.⁴³ However, an inverted funnel plot showed no evidence of publication bias. Finally, the 13 included studies used a variety of reference tests (cervical mediastinoscopy, anterior mediastinotomy and/or thoracotomy with ipsilateral lymph node sampling), raising the possibility of differential verification

bias.⁴¹ Because none of the reference tests has perfect sensitivity, the true sensitivity of TBNA may be even lower than our estimates. Future studies of the diagnostic accuracy of TBNA should require thoracotomy with systematic sampling of both normal and abnormal appearing lymph nodes at all accessible mediastinal stations to exclude the presence of lymph node metastasis.⁴⁴

In conclusion, we found that TBNA is highly specific for detecting mediastinal lymph node metastasis in patients with NSCLC, but that sensitivity depends critically on the prevalence of mediastinal lymph node involvement. In patient populations with a relatively low prevalence of mediastinal disease (such as those with potentially resectable NSCLC), the sensitivity of TBNA is poor.

ACKNOWLEDGEMENTS

The authors thank Helena C Kraemer for providing statistical advice; Corinna Haberland, Annette Langer-Gould, Pavel Strnad, Kelvin Tan, and Tomasz Ziedalski for reviewing non-English language studies; and the authors of the studies included in our meta-analysis who provided additional unpublished data.



A more detailed description of the methods used in this study is available as an online supplement on the Thorax website at www.thoraxjnl.com/supplemental

Authors' affiliations

J-E C Holty, W G Kuschner, M K Gould, VA Palo Alto Health Care System, Palo Alto, CA and Stanford University School of Medicine, Stanford, CA, USA

J-E C Holty, M K Gould, Center for Primary Care and Outcomes Research, Stanford University, Stanford, CA, USA

Dr Holty is supported by the VA Special Fellowships Program. Dr Gould is a recipient of an Advanced Research Career Development Award from the VA Health Services Research and Development Service.

Drs Holty, Gould and Kuschner have no financial conflict of interest or competing interests to disclose.

REFERENCES

- 1 **Spiro SG**, Porter JC. Lung cancer - where are we today? Current advances in staging and nonsurgical treatment. *Am J Respir Crit Care Med* 2002;**166**:1166-96.
- 2 **Ries LAG**. Influence of extent of disease, histology and demographic factors on lung cancer survival in the SEER population-based data. *Semin Surg Oncol* 1994;**10**:21-30.
- 3 **Mountain CF**. Revisions in the international system for staging lung cancer. *Chest* 1997;**111**:1710-7.
- 4 **Patterson GA**. Surgical techniques in the diagnosis of lung cancer. *Chest* 1991;**100**:523-6.
- 5 **Arita T**, Matsumoto T, Kuramitsu T, *et al*. Is it possible to differentiate malignant mediastinal nodes from benign nodes by size? Reevaluation by CT, transesophageal echocardiography and nodal specimen. *Chest* 1996;**110**:1004-8.
- 6 **Takeshi A**, Tatsuya K, Mitsutoshi M. Bronchogenic carcinoma: incidence of metastases to normal sized lymph nodes. *Thorax* 1995;**50**:1267-9.
- 7 **Lahde S**, Paivansalo M, Rainio P. CT for predicting the resectability of lung cancer. A prospective study. *Acta Radiol* 1991;**32**:449-54.
- 8 **Dales RE**, Stark RM, Raman S. Computed tomography to stage lung cancer. Approaching a controversy using meta-analysis. *Am Rev Respir Dis* 1990;**141**:1096-101.
- 9 **Gould MK**, Kuschner WG, Rydzak CE, *et al*. Test performance of positron emission tomography and computed tomography for mediastinal staging in patients with non-small-cell lung cancer. *Ann Intern Med* 2003;**139**:879-92.
- 10 **Rami Porta R**. Surgical exploration of the mediastinum by mediastinoscopy, parasternal mediastinotomy and repleurostomy: indications, technique and complications. *Ann Ital Chir* 1999;**70**:867-72.
- 11 **Hujala KT**, Sipila JI, Grenman R. Mediastinoscopy: its role and value today in the differential diagnosis of mediastinal pathology. *Acta Oncol* 2001;**40**:79-82.

- 12 **Detterbeck FC**, DeCamp MM Jr, Kohman LJ, *et al*. Lung cancer. Invasive staging: the guidelines. *Chest* 2003;**123**(1 Suppl):167-75S.
- 13 **Shure D**. Fiberoptic bronchoscopy: diagnostic applications. *Clin Chest Med* 1987;**8**:1-13.
- 14 **Tolosa EM**, Harpole L, Detterbeck F, *et al*. Invasive staging of non-small cell lung cancer: a review of the current evidence. *Chest* 2003;**123**(1 Suppl):157-66S.
- 15 **Kent DL**, Haynor DR, Larson EB, *et al*. Diagnosis of lumbar spinal stenosis in adults: a meta-analysis of the accuracy of CT, MR, and myelography. *AJR Am J Roentgenol* 1992;**158**:1135-44.
- 16 **Gould MK**, Maclean CC, Kuschner WG, *et al*. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions. A meta-analysis. *JAMA* 2001;**285**:914-24.
- 17 **Harkin TJ**, Wang KP. Bronchoscopic needle aspiration of mediastinal and hilar lymph nodes. *J Bronchology* 1997;**4**:238-49.
- 18 **Dasgupta A**, Mehta AC, Wang KP. Transbronchial needle aspiration. *Semin Respir Crit Care Med* 1997;**18**:571-81.
- 19 **Harrow EM**, Wang KP. The staging of lung cancer by bronchoscopic transbronchial needle aspiration. *Chest Surg Clin North Am* 1996;**6**:223-35.
- 20 **Kraemer HC**, Periyakoil VS, Noda A. Tutorial in biostatistics: kappa coefficients in medical research. *Stat Med* 2002;**21**:2109-29.
- 21 **Thompson JR**. Estimating equations for kappa statistics. *Stat Med* 2001;**20**:2895-906.
- 22 **DerSimonian R**, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;**7**:177-88.
- 23 **Moses LE**, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med* 1993;**12**:1293-316.
- 24 **Hedges LV**, Olkin I. Regression models in research synthesis. *Am Stat* 1983;**37**:137-40.
- 25 **Olkin I**, Sampson A. Comparison of meta-analysis versus analysis of variance of individual patient data. *Biometrics* 1998;**54**:317-22.
- 26 **Cooper H**, Hedges LV. *The Handbook of research synthesis*. New York: Russell Sage Foundation, 1994.
- 27 **Fleiss JL**. *Statistical methods for rates and proportions*, 2nd edn. New York: John Wiley & Sons, 1981.
- 28 **Harrow EM**, Abi-Saleh W, Blum J, *et al*. The utility of transbronchial needle aspiration in the staging of bronchogenic carcinoma. *Am J Respir Crit Care Med* 2000;**161**:601-7.
- 29 **Bilaceroglu S**, Cagirci U, Gunel O, *et al*. Comparison of rigid and flexible transbronchial needle aspiration in the staging of bronchogenic carcinoma. *Respiration* 1998;**65**:441-9.
- 30 **Disdier C**, Varela G, Sanchez de Cos J, *et al*. Utilidad de la puncion tranbronquial y la mediastinoscopia en la estadiacion ganglionar mediastinica del carcinoma broncogenico no microcitico. Estudio preliminar. *Arch Bronconeumol* 1998;**34**:237-44.
- 31 **Raito GB**, Mereu C, Motta G. The prognostic significance of preoperative assessment of mediastinal lymph nodes in patients with lung cancer. *Chest* 1988;**93**:807-13.
- 32 **Schenk DA**, Bower JH, Bryan CL, *et al*. Transbronchial needle aspiration staging of bronchogenic carcinoma. *Am Rev Respir Dis* 1986;**134**:146-8.
- 33 **Herth F**, Hecker E, Hoffman H, *et al*. [Endobronchial ultrasound for local tumour and lymph node staging in patients with centrally growing lung cancer]. *Ultraschall Med* 2002;**23**:251-66.
- 34 **Wang R**, Han JJ, Yao J, *et al*. [The role of transbronchial needle aspiration in the staging of bronchogenic carcinoma]. *Zhongguo fei ai za zhi (Chinese Journal of Lung Cancer)* 2002;**5**:284-6.
- 35 **Patelli M**, Lazzari L, Poletti V, *et al*. Role of fiberoptic transbronchial needle aspiration in the staging of N2 disease due to non-small cell lung cancer. *Ann Thorac Surg* 2002;**73**:407-11.
- 36 **Katis K**, Kotrogiannis G, Paraskevopoulos I, *et al*. Bronchoscopic needle aspiration in mediastinal staging of patients with bronchogenic carcinoma. *J Bronchology* 1998;**5**:195-9.
- 37 **Rong F**, Cui B. CT scan directed transbronchial needle aspiration biopsy for mediastinal nodes. *Chest* 1998;**114**:36-9.
- 38 **Schenk DA**, Chambers SL, Derdak S, *et al*. Comparison of the Wang 19-gauge and 22-gauge needles in the mediastinal staging of lung cancer. *Am Rev Respir Dis* 1993;**147**:1251-8.
- 39 **Schenk DA**, Strollo PJ, Pickard JS, *et al*. Utility of the Wang 18-gauge transbronchial histology needle in the staging of bronchogenic carcinoma. *Chest* 1989;**96**:272-4.
- 40 **Wang KP**, Brower R, Haponik EF, *et al*. Flexible transbronchial needle aspiration for staging of bronchogenic carcinoma. *Chest* 1983;**84**:571-6.
- 41 **Whiting P**, Rutjes AWS, Reitsma JB, *et al*. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 2004;**140**:189-202.
- 42 **Brenner H**, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat Med* 1997;**16**:981-91.
- 43 **Wiersema LM**, Edell ES, Midhun DE, *et al*. A blinded prospective study comparing transbronchial needle aspiration (TBNA) and endosonography guided needle aspiration (EUS FNA) in the staging of mediastinal lymphadenopathy (MLN) in patients with non small cell lung cancer (NSCLC). *Gastrointest Endosc* 2001;**53**:AB68.
- 44 **Gould MK**, Silvestri GA, Detterbeck F. Multidisciplinary management of lung cancer. *N Engl J Med* 2004;**350**:2008-10.