National Institute for Health and Care Excellence

Guideline version (Draft)

Lung Cancer Update

Evidence reviews for effectiveness of non-ultrasound-guided TBNA, EBUS-TBNA or EUS-FNA for people with a probability of mediastinal malignancy

NICE guideline <number>

Evidence reviews

October 2018

Draft for Consultation

These evidence reviews were developed by the NICE Guideline Updates Team

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

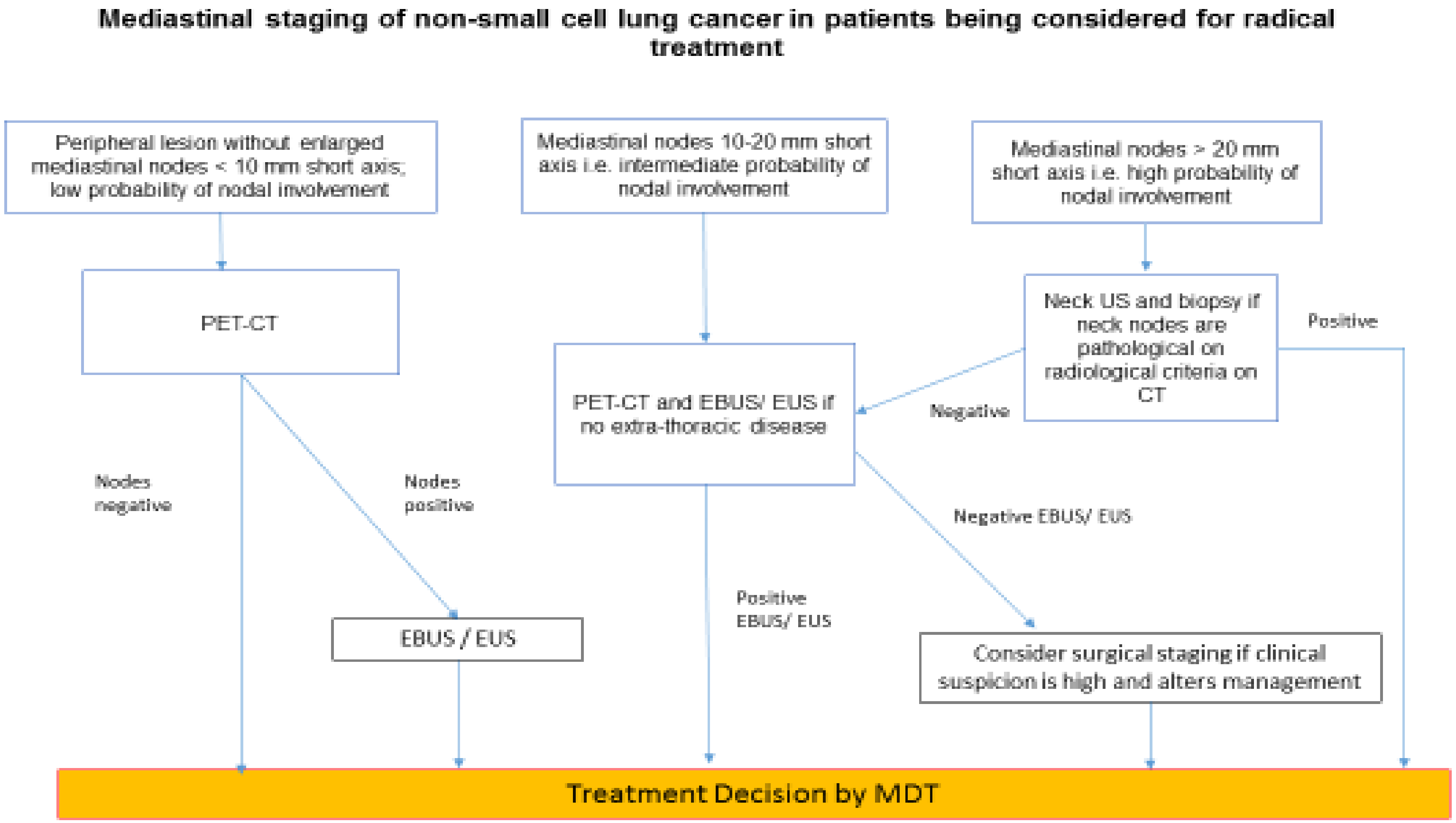
NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN:

Contents

Evidence reviews for clinical and cost effectiveness of non-ultrasound-guided TBNA, EBUS-TBNA or EUS-FNA alone or in combination for people with a probability of mediastinal malignancy


Mediastinal staging of non-small cell lung cancer in patients being considered for radical treatment
Peripheral lesion without enlarged mediastinal nodes < 10 mm short axis; low probability of nodal involvement
Mediastinal nodes 10-20 mm short axis i.e. intermediate probability of nodal involvement
Mediastinal nodes > 20 mm short axis i.e. high probability of nodal involvement
PET-CT
Neck US and biopsy if neck nodes are pathological on radiological criteria on CT
Positive
PET-CT and EBUS/ EUS if no extra-thoracic disease
Negative
Nodes negative
Nodes positive
Negative EBUS/ EUS
Positive EBUS/ EUS
EBUS / EUS
Consider surgical staging if clinical suspicion is high and alters management
Treatment Decision by MDT

1
2

Review questions

RQ 1.1: What is the clinical and cost effectiveness of using non-ultrasound-guided TBNA, EBUS-TBNA or EUS-FNA as the first invasive test for people with a probability of mediastinal malignancy?

RQ 1.2: What is the clinical and cost-effectiveness of EBUS-TBNA alone, EUS-FNA alone or EBUS-TBNA and EUS-FNA in combination compared with surgical staging to diagnose and/or stage lung cancer?

Introduction

Since publication of the existing guideline CG121, a randomised controlled trial (RCT) suggested that the use of endobronchial ultrasound transbronchial needle aspiration (EBUS-TBNA) and occasional use of endoscopic ultrasound-guided fine needle aspiration (EUS-FNA) in the diagnosis of lung cancer enabled:

- faster treatment decisions compared to conventional diagnosis and staging;
- fewer invasive investigations per person compared to conventional diagnosis and staging;
- improved survival (all-cause hazard ratio) compared to conventional diagnosis and staging in a post-hoc analysis (Navani 2015).

Conventional diagnosis and staging included CT-guided biopsy and non-ultrasound-guided TBNA. Another RCT suggested that EBUS-TBNA in combination with EUS-FNA is more effective and less expensive than standard surgical staging alone (Annema 2010, Sharples 2012). Therefore, the purposes of this review are to:

- Determine the effectiveness of using non-ultrasound-guided TBNA, EBUS-TBNA or EUS-FNA as the first invasive test for people with a probability of mediastinal malignancy.
- Determine the effectiveness of EBUS-TBNA alone, EUS-FNA alone or EBUS-TBNA and EUS-FNA in combination compared with surgical staging to diagnose and/or stage lung cancer.

Table 1: PICO table

Population	Patients with suspected/ confirmed lung cancer (Pre-diagnosis and CT std. clinical evaluation)
Interventions	• Non-ultrasound-guided TBNA, • EBUS-TBNA or • EUS-FNA
Comparator	The gold standard investigation (histological/ cytological confirmation and pathological TNM - Or follow up period adequate to confirm outcome - Normally pathology from surgical resection but could be another technique in specified circumstances.
Outcomes	• The diagnostic sensitivity and specificity (likelihood ratios) • The staging sensitivity and specificity • The safety of each procedure/ adverse events (EBUS – mortality, in-patient admission, pneumothorax) • Patient acceptability • Anxiety and psychological outcomes – report if in evidence • Timing (for example, time to treatment) • The number of investigations and outpatient attendances per patient

Table 2 PICO table

Population	Patients with suspected/ confirmed lung cancer (Pre-diagnosis and CT std. clinical evaluation)
Interventions	• EBUS-TBNA alone, • EUS-FNA alone or • EBUS-TBNA and EUS-FNA in combination
Comparator	• Surgical staging Or follow up period adequate to confirm outcome - Normally pathology from surgical resection but could be another technique in specified circumstances.
Outcomes	• The diagnostic sensitivity and specificity (likelihood ratios) • The staging sensitivity and specificity • The safety of each procedure/ adverse events (EBUS – mortality, in-patient admission, pneumothorax) • Patient acceptability • Anxiety and psychological outcomes – report if in evidence • Quality of life • The number of investigations and outpatient attendances per patient • Timing (for example, time to treatment)

Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in appendix A, and the methods section in appendix B. In particular, the minimally important differences (MIDs) used in this review are summarised in appendix B.

Declarations of interest were recorded according to NICE's 2018 conflicts of interest policy.

Clinical evidence

Included studies

This review was conducted as part of a larger update of the NICE Lung cancer: diagnosis and management guideline (CG121). A systematic literature search for RCTs and systematic reviews with no date limit yielded 2,117 references.

Papers returned by the literature search were screened on title and abstract, with 48 full-text papers ordered as potentially relevant systematic reviews or RCTs. RCTs were excluded if they did not meet the criteria of enrolling patients with suspected or confirmed lung cancer.

Six papers representing 5 unique RCTs were included after full text screening. Three of these were cross-sectional diagnostic RCTs: Annema 2010 (n=241, follow-up period 1 year), Kang 2014 (n=160, follow-up period 3-5 days), Tournoy 2008 (n=40 days, median follow-up period 2 nights). Two studies were interventional RCTs: Larsen 2005 (n=104, median follow-up period 1.3 and 1.4 years for each arm respectively) and Navani 2015 (n=132, median follow-up period 503 days and 312 days for each arm respectively). Multiple papers reporting results of the same study were identified and collated, so that each study rather than individual reports was the unit of interest in the review, therefore there were 5 unique studies. The following reference standards were used - for benign results: surgical confirmation and for malignant results: pathology.

For the search strategy, please see appendix C. For the clinical evidence study selection flowchart, see appendix D. For the full evidence tables and full GRADE profiles for included studies, please see appendix E and appendix F.

Excluded studies

Details of the studies excluded at full-text review are given in appendix G along with a reason for their exclusion.

Summary of clinical studies included in the evidence review

Five randomised controlled studies were included in this review. The following studies met the inclusion criteria for RQ 1.1: Larsen 2005 and Tournoy 2008. The following study met the inclusion criteria for RQ 1.2: Annema 2010. The following studies met the inclusion criteria for both RQ 1.1 and 1.2: Kang 2014 and Navani 2015.

Study locations

One randomised controlled study was from the UK (Navani 2015), 1 was from the Netherlands, Belgium and the UK (Annema 2010), 1 was from South Korea (Kang 2014), 1 was from Denmark (Larsen 2005) and 1 was from Belgium (Tournoy 2008).

Outcomes and sample sizes

The reported outcomes with extractable data were diagnostic performance (preferably sensitivity, diagnostic negative predictive value, staging sensitivity), mortality, in-patient admission, pneumothorax, other complications, patient acceptability, anxiety and psychological problems, time to treatment decision, time to diagnosis and staging, number of investigations per person, number of outpatient attendances per person and quality of life. Additional non-protocol outcome measures were recorded. Rather than exclude them, the committee decided that they were worthy of consideration. The non-protocol outcome measures were: number of avoidable thoracotomies and recurrence during a specified follow-up time. The committee wanted to know the number of avoidable thoracotomies because unnecessary thoracotomies can be distressing for patients. Recurrence during a specified follow-up time was useful for the economic modelling. The sample sizes ranged from 40 participants to 257 across studies.

See full evidence tables and GRADE profiles Appendix E and Appendix F.

Quality assessment of clinical studies included in the evidence review

See appendix F for full GRADE tables.

Economic evidence

Standard health economic filters were applied to the clinical search for this question, and a total of 1,788 citations was returned. Details of the literature search are provided in Appendix C. Following review of titles and abstracts, 24 full-text studies were retrieved for detailed consideration. One relevant cost–utility analysis, 1 health economics paper with a survival model and one health economics paper with an influence diagram were identified. Therefore 3 studies were included in this review.

EBUS-FNA plus EBUS-TBNA vs surgical staging

Sharples et al. (2012) conducted a cost-utility analysis alongside a 6-month RCT (ASTER) in the UK, Belgium and the Netherlands (n=247). Patients were eligible for the trial if they had known/suspected non-small cell lung cancer (NSCLC), with suspected mediastinal lymph node involvement; otherwise eligible for surgery with curative intent; clinically fit for endosonography and surgery; and had no evidence of metastatic disease. Patients were excluded from the trial if they had previous lung cancer treatment; concurrent malignancy; uncorrected coagulopathy; or were not suitable for surgical staging. One hundred and twenty three patients were randomised to endosonography followed by surgical staging if no nodal metastases were found at endosonography, whilst 118 patients were randomised to surgical staging alone. The primary research objective of the study was to determine whether endosonography is better than standard surgical staging techniques in terms of sensitivity, diagnostic accuracy and negative predictive value for diagnosing and staging the mediastinum in lung cancer. A secondary research objective was to conduct a comparative cost analysis of the diagnostic strategies of the two trial arms.

Endosonography in this study was EBUS-TBNA combined with EUS-FNA. Surgical staging was performed by (video) mediastinoscopy, left anterior mediastinoscopy or video-assisted thoracoscopy or combination.

The authors' base case adopted a UK NHS perspective. Resource use was collected in terms of numbers of procedures done, (surgical, radiotherapy, chemotherapy) treatments administered, hospital and hospice stays. Costs were taken from the Department of Health (DoH) NHS reference costs 2008-2009. Cost estimates for endosonography were estimated by Papworth Hospital finance department. The price year was 2008-2009.

Utility was measured using the EQ-5D at baseline, end of staging, 2 months and 6 months, using a UK tariff.

Bayesian parametric modelling was used to estimate final expected costs and quality-adjusted life-years (QALYs) while simultaneously estimating missing data based on randomisation group, centre and stage.

Base-case results for patients for whom complete information on trial costs and QALYs were available (endosonography n=58, surgical staging n=56) are shown in Table 3.

Table 3: Costs and effects from Sharples et al. (2012)

	Absolute		Incremental		
Strategy	**Cost**	**Effect**	**Cost**	**Effect**	**ICER**
Surgical Staging Alone	11,735 £GBP (10,843 to 12,647)	0.342 QALYs (0.316 to 0.367)			
Endosonography followed by Surgical Staging	10,808 £GBP (9,843 to 11,764)	0.348 QALYs (0.321 to 0.373)	-927 £GBP (-2246 to 394)	0.00652 QALYs (-0.0298 to 0.0418)	Endosonography followed by Surgical Staging **Dominant**

Endosonography followed by surgical staging compared to surgical staging alone was £972 cheaper and produced 0.00652 more QALYs, rending endosonography followed by surgical staging as a dominant strategy. (Strategies that are dominant cost less and are more effective than their comparator.)

Because of the very small QALY difference, the authors concluded that an ICER could not be reliably estimated but in the probabilistic sensitivity analysis, 63% of bootstrapped samples showed endosonography dominated (which means it was less expensive and produced more benefit compared to) surgical staging and endosonography was cost-effective at a threshold of £30,000/QALY in 99.9% of samples.

EBUS-TBNA vs conventional approaches

Navani et al. (2015) conducted a cost-effectiveness analysis alongside LUNG-BOOST, an open-label, multicentre, pragmatic, randomised controlled trial. Patients were recruited from 6 centres in the UK, who were suspected to have stage I to IIIA lung cancer on the basis of CT scans of the neck, thorax, and upper abdomen were eligible for trial entry. For inclusion into the trial, patients had to be aged at least 18 years and fit enough to undergo thoracotomy and lung resection. Exclusion criteria were significant concurrent malignant disease or any condition or concurrent medicine that contraindicated EBUS-TBNA or mediastinoscopy. Patients with known extrathoracic malignant disease, supraclavicular lymphadenopathy, or pleural effusion were also excluded. Of the 133 RCT participants, 66 participants were randomised to endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA), whilst 67 patients were randomised to conventional diagnosis and staging (CDS).

The primary endpoint was the time from first outpatient appointment with the respiratory specialist to treatment decision by the multidisciplinary team, after completion of the diagnosis and staging procedures. Analysis took a UK NHS perspective.

Effectiveness in this study was measured using mean time to treatment decision from the first outpatient appointment with the respiratory specialist, using hazard ratios. This is in contrast to the NICE reference case, where effects are measured in QALYs. Unit costs were obtained from NHS reference costs, NICE 2011 lung cancer guideline, and a published study; these were multiplied by the resource use and summed across all resource items. The price year was 2010-2011.

Lung cancer was diagnosed in 57 (86%) patients in the CDS group and 50 (76%) in the EBUS group ($p=0{\cdot}196$), and clinical staging did not differ significantly between the groups in patients with non-small-cell lung cancer.

The median time-to-treatment decision was longer after CDS (29 days [95% CI 23–35]), than after EBUS (14 days [14–15]; HR 1·98, 95% CI 1·39–2·82, $p<0{\cdot}0001$) in the intention-to-diagnose population. Therefore, patients in the EBUS group of the trial were likely to receive a treatment decision twice as fast as patients in the CDS group. A greater proportion of patients had diagnosis and staging completed by 14 days in the EBUS group than in the CDS group (35 [53%] vs 8 [12%], $p<0{\cdot}0001$). In the subset of patients with non-small-cell lung cancer, initial EBUS-TBNA resulted in a shorter time-to-treatment decision of 15 days (95% CI 14–16), compared with 30 days (95% CI 23–34) in the CDS group (HR 2·09, 95% CI 1·38–3·15, $p=0{\cdot}0002$).

In a post-hoc analysis, the median survival of patients with non-small-cell lung cancer in the EBUS group of 503 days (95% CI 312–715) was longer than the median survival in the CDS group of 312 days (95% CI 231–488; HR 0·60, 0·37–0·98, p=0·0382;). An exploratory analysis of lung cancer patients who underwent surgery suggested that postoperative survival was better in the EBUS group than in the CDS group.

For diagnosis and staging, EBUS-TBNA was found to cost £2,407 (SD £180.50) whilst CDS was found to cost £2,348 (SD £192.20). This represents an incremental cost for EBUS-TBNA of £59 (95% CI –£463 to £581). Mean initial treatment costs per patient in those diagnosed with lung cancer were £4452 (£180·00) and £4261 (£257·90), respectively (difference £191, 95% CI –447 to 829).

The results from the trial suggest that routine use of EBUS-TBNA as an initial investigation after a staging CT for suspected lung cancer scan results in a faster treatment decision, with fewer investigations at no significant difference in cost, and, in post-hoc analysis, seems to improve survival, compared with conventional diagnosis and staging methods.

Influence Diagram model to determine optimal sequence of tests for mediastinal staging of lung cancer

Luque et al. (2016) created an influence diagram (ID) model for a Spanish public healthcare system to determine the optimal sequence of tests for the mediastinal staging of non-small cell lung cancer (NSCLC) by considering sensitivity, specificity, and the economic cost of each test. This was stated to be important, as correct staging of the disease as early as possible helps to determine which patients may benefit from surgery and, in turn, to avoid dangerous, painful, and unnecessary surgery when metastasis has already occurred.

The model assumed that all patients first had a computed tomography (CT) scan, and then could have a transbronchial needle aspiration (TBNA), positron emission tomography (PET), endobronchial ultrasound (EBUS), endoscopic ultrasound (EUS), or a mediastinoscopy (MED) in various sequences.

IDs are a new modelling method that makes use of advanced statistical and computer science techniques to handle problems where the numbers of sequential decisions and probabilities are too large to be easily evaluated by a conventional decision tree. An auxiliary Bayesian network was built that could handle every possible sequence of tests as well as patients' decisions and outcomes.

The ID model was evaluated twice, first without considering economic costs, and then considering cost effectiveness using a willingness-to-pay of €30,000 per QALY, the shadow threshold estimated for the Spanish health system. The authors performed several types of sensitivity analysis to study the effect of the uncertainty in the numerical parameters of the model.

The authors reported the optimal strategies using the two different criteria. When considering only effectiveness, a positive computed tomography (CT) scan should be followed by a transbronchial needle aspiration (TBNA) and an endobronchial ultrasound (EBUS). Endoscopic ultrasound (EUS) and mediastinoscopy are then used to either confirm negative findings or when the results of two tests are contradictory. When the CT scan is negative, a positron emission tomography (PET) and EBUS are performed. EUS and mediastinoscopy are used in the case of negative or contradictory results.

Economic model conducted for the 2011 NICE lung cancer guideline

The economic model built for the 2011 NICE lung cancer guideline included a range of diagnosis and staging strategies for people with an intermediate probability of mediastinal malignancy.

The model was a decision tree comprising 27 possible strategies which included one or several of neck ultrasound, PET-CT, conventional TBNA, EBUS TBNA and mediastinoscopy in various orders. Patients at each final end point entered a two state Markov model comprising survival and death states.

Disease prevalence, distribution of treatment options and survival estimates were drawn from registry data and expert opinion. Costs were drawn from standard NHS sources and resource use was drawn from expert opinion. The test accuracy data was drawn from expert opinion. Utility data were drawn from published literature and expert opinion.

The model concluded that PET-CT followed by conventional TBNA was the optimal strategy. This was due to the combination of high sensitivity and low cost parameters used within the model for these tests. The model was reasonably robust with regards to deterministic sensitivity analysis but no probabilistic sensitivity analysis was conducted. The guideline committee concluded that while the model had a number of limitations, the results provided them with useful information when developing a diagnostic testing algorithm.

Evidence statements

EUS-FNA followed by EBUS-TBNA vs straight to surgical staging

Effectiveness data

Low to moderate-quality evidence from 1 RCT reporting data on 241 people with suspected N2 or N3 mediastinal lymph node involvement found that there was a greater number of avoidable thoracotomies in people offered EUS-FNA followed by EBUS-TBNA compared to people who went straight to surgical staging. However, there was no difference in the number of people experiencing a pneumothorax, the total number of complications, quality of life at 6 months, or the number of people who died between staging and 6 months later.

Diagnostic accuracy data

Moderate-quality evidence from 1 RCT reporting data on 241 people with suspected N2 or N3 mediastinal lymph node involvement found the sensitivity of EUS-FNA followed by EBUS-TBNA was 93.3% and the negative predictive value was 92.7% (with a prevalence of 53.7%). The sensitivity of the straight to surgical staging arm was 78.3% and the negative predictive value was 85.3% (with a prevalence of 44.1%).

Bronchoscopy, EBUS-TBNA then EUS-FNA if necessary vs bronchoscopy, EUS-FNA then EBUS-TBNA if necessary

Effectiveness data

Very low-quality evidence from 1 RCT reporting data from 160 people with histologically confirmed or strongly suspected, potentially operable non-small cell lung cancer found that the data could not differentiate the number of people experiencing a pneumothorax or patient tolerance 3-5 days after the interventions.

Diagnostic accuracy data

Low-quality evidence from 1 RCT reporting data from 148 people with histologically confirmed or strongly suspected, potentially operable non-small cell lung cancer found that the sensitivity of bronchoscopy, EBUS-TBNA, then EUS-FNA was 85.3% and the negative predictive value was 88.0% (with a prevalence of 45.9%). The sensitivity of bronchoscopy, EUS-FNA, then EBUS-TBNA was 90.4% and the negative predictive value was 95.2% (with a prevalence of 33.8%). For the bronchoscopy, EBUS-TBNA, then EUS-FNA arm, the sensitivity of EBUS-TBNA was 81.4% and its negative predictive value was 86.2% (with a prevalence 45.9%). In the bronchoscopy, EUS-FNA, then EBUS-TBNA arm, the sensitivity of EUS-FNA was 59.6% and its negative predictive value was 82.5% (with a prevalence of 33.8%).

Mediastinoscopy + EUS-FNA vs mediastinoscopy + EUS-FNA only if CT shows invasion adjacent to the oesophagus

Effectiveness data

High-quality evidence from 1 RCT reporting data from 104 people with suspected or diagnosed lung cancer after CT/PET, bronchoscopy, TBNA/TTNA, lung function tests and general examination found that there was a greater number of avoidable thoracotomies in the mediastinoscopy + EUS-FNA arm compared to the mediastinoscopy + EUS-FNA only if CT shows invasion adjacent to the oesophagus arm. However, moderate-quality data could not differentiate between complications, recurrence or death.

EBUS-TBNA (or EUS-FNA) vs conventional diagnosis and staging (bronchoscopy or CT-guided biopsy etc.)

Effectiveness data

High to moderate-quality evidence from 1 RCT reporting data from 132 people with suspected stage I to IIIA lung cancer on CT neck, thorax and upper abdomen showed that there was a reduction in time to treatment decision, a reduction in the number of investigations per person, an increase in the duration of survival (hazard ratio), an increase in the number of people who had diagnosis and staging competed by 14 days and an increase in the number of people diagnosed and staged with one investigation for EBUS-TBNA (or EUS-FNA) compared to conventional diagnosis and staging (bronchoscopy or CT-guided biopsy etc.) However, the data could not differentiate between the number of avoidable thoracotomies and the number of people experiencing a pneumothorax or in-patient admissions.

Diagnostic accuracy data

High to moderate-quality evidence from 1 RCT reporting data from 132 people with suspected stage I to IIIA lung cancer on CT neck, thorax and upper abdomen showed that for EBUS-TBNA (or EUS-FNA) the sensitivity was 92.0% and the negative predictive value was 90.0% (with a prevalence of 75.8%).

EUS-FNA vs straight to surgical staging

Effectiveness data

Moderate to low-quality evidence from 1 RCT reporting data from 40 people who had proven or suspected NSCLC or suspected mediastinal lymph node invasion on CT/PET found that the date could not differentiate the numbers of people experiencing perforation or bleeding.

Diagnostic accuracy data

Low-quality evidence from 1 RCT reporting data from 40 people who had proven or suspected NSCLC or suspected mediastinal lymph node invasion on CT/PET found that the sensitivity for EUS-FNA for all was 93.0% and the negative predictive value was 83.0% (with a prevalence of 73.7%). For people who went straight to surgical staging, the sensitivity was 73.0% and the negative predictive value was 73.0% (with a prevalence of 52.3%).

Reference standards: For benign results, surgical confirmation. For malignant results, pathology.

Health economics evidence statements

One directly applicable UK, Belgian and Dutch based cost-utility analysis with potentially serious limitations compared endosonography followed by surgical staging with surgical staging alone for the staging of potentially resectable lung cancer. Endosonography followed by surgical staging compared to surgical staging alone was found to be a dominant strategy. A cost-effectiveness acceptability curve (CEAC) for endosonography followed by surgery if negative showed that 92% of the scenarios involved cost savings.

One partially applicable UK cost-effectiveness analysis with potentially serious limitations compared endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA), to conventional diagnosis and staging (CDS) for diagnosis and staging in patients who were suspected to have stage I to IIIA lung cancer on the basis of CT scans of the neck, thorax, and upper abdomen. EBUS-TBNA for investigation was found to be slightly more expensive than CDS, but resulted in a shortened median time to treatment decision of nearly 50%. A post-hoc analysis revealed that the median survival time was greater for those in the EBUS-TBNA arm of the trial compared to those in the CDS arm.

One directly applicable economic model with very serious limitations found that PET-CT followed by conventional TBNA was the most cost effective strategy for people with an intermediate probability of mediastinal malignancy.

One partially applicable influence diagram with very serious limitations found that when considering only effectiveness, the optimal strategy following a positive computed tomography (CT) scan was transbronchial needle aspiration (TBNA), followed by an endobronchial ultrasound (EBUS), and an endoscopic ultrasound (EUS). When the CT scan is negative, the optimal strategy was positron emission tomography (PET) followed by EBUS, and EUS. When taking into account costs, the optimal strategy following a positive CT scan was TBNA only; with an EBUS being done only when the CT scan or the TBNA is negative.

Recommendations

Effectiveness of diagnostic and staging investigations

1.3.10 Audit the local test performance of EBUS-TBNA and endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA). [2011, amended 2019]

1.3.11 When taking samples, ensure they are adequate (without unacceptable risk to the person) to permit pathological diagnosis, including tumour subtyping and assessment of predictive markers. [2011, amended 2019]

Sequence of investigations

1.3.13 Choose investigations that give the most information about diagnosis and staging with least risk to the person. Think carefully before performing a test that gives only diagnostic pathology when information on staging is also needed to guide treatment. [2011]

1.3.14 Perform CT of chest, liver, adrenal and lower neck[1] before:

- an intended bronchoscopy or EBUS
- any other biopsy procedure. [2005, amended 2019]

Peripheral primary tumour

1.3.15 Offer image-guided biopsy to people with peripheral lung lesions when treatment can be planned on the basis of this test. [2011, amended 2019]

1.3.16 Biopsy any enlarged mediastinal nodes (10 mm or larger maximum short axis on CT) or other lesions in preference to the primary lesion if determination of stage affects treatment2 [2011]

Central primary tumour

1.3.17 Offer flexible bronchoscopy to people with central lesions on CT if nodal staging does not influence treatment. [2011, amended 2019]

Mediastinal lymph node assessment

1.3.18 Offer PET-CT as the preferred first test after CT with a low probability of mediastinal malignancy (lymph nodes below 10 mm maximum short axis on CT), for people with lung cancer who could potentially have treatment with curative intent. [2011]

1.3.19 Offer PET-CT (if not already done), and one or both of EBUS-TBNA and EUS-FNA, as the initial investigations for people with lung cancer who have an intermediate probability of mediastinal malignancy (lymph nodes between 10 and 20 mm maximum short axis on CT) and who could potentially have treatment with curative intent. [2019]

1.3.20 Offer neck ultrasound with sampling of visible lymph nodes to people with a high probability of mediastinal malignancy (lymph nodes over 20 mm maximum short axis on CT). If neck ultrasound is negative, follow with EBUS-TBNA and/or EUS-FNA. [2019]

1.3.21 Evaluate PET-CT-positive mediastinal nodes with EBUS-TBNA and/or EUS-FNA if nodal status would affect the treatment plan. [2019]

1.3.22 Consider surgical mediastinal staging for people with a negative EBUS-TBNA or EUS-FNA if clinical suspicion of mediastinal malignancy is high and nodal status would affect their treatment plan. [2019]

---

[1] This recommendation was outside the scope of the 2019 update, but the guideline committee recognised that many centres include the lower neck when performing CT scans before bronchoscopy, EBUS and other biopsy procedures. The committees also recognised that contrast medium should only be given with caution to people with known renal impairment.

[2] Some people with lung cancer will not be well enough for treatment with curative intent. This needs to be taken into account when choosing diagnostic and staging investigations

Rationale and impact

Why the committee made the recommendations: Effectiveness of diagnostic and staging investigations

Recommendation 1.3.10

Clinical audit is an important tool for maintaining high standards in the use of endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA) and endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA).This is consistent with the British Thoracic Society guideline and quality standards (which are endorsed by NICE).

Why the committee made the recommendations: EBUS-TNBA and EUS-FNA

Recommendations 1.3.19-1.3.21

The recommendations cover:

- initial invasive investigations for people with an intermediate probability of mediastinal malignancy
- subsequent investigations for people with a high probability of mediastinal malignancy, when neck ultrasound and biopsy are negative.

In these circumstances, when compared with alternative investigations EBUS-TBNA and EUS-FNA:

- produce a diagnosis faster than bronchoscopy or CT-guided biopsy
- are more acceptable to patients than surgery
- reduce the need for further endoscopic investigations and hospital visits compared with bronchoscopy.

The decision on which procedure to use depends on where suspicious lesions are located. For example, EUS-FNA enters the area between the lungs through the oesophagus, so can more easily access lung stations 8, 9 and 4L. By contrast, EBUS-TBNA enters the area between the lungs through the trachea, so can more easily access lung stations closer to the large airways. The recommendations do not specify when one procedure is better than the other because there is variation in the way that suspicious lesions show up on imaging, so evidence was not available for every possible situation. Because of this, clinicians will need to use their judgement on whether to use EBUS-TBNA, EUS-FNA, or both.

The availability of PET-CT is more limited than EBUS-TBNA and EUS-FNA, so specifying that PET-CT is done first may cause delays in diagnosis. As a result, the committee did not recommend a specific order for the investigations.

Why the committee made the recommendations: Surgical mediastinal staging

Recommendation 1.3.22

When EBUS-TBNA and/or EUS-FNA are negative but clinical suspicion of mediastinal malignancy is high, surgical mediastinal staging is the final staging option. Nodal status may affect the treatment plan. While there are potential harms from the invasive nature of surgical staging, there is no evidence that these outweigh the benefits in this population. With these points in mind, the committee recommended consideration of surgical mediastinal staging based on their knowledge and experience.

Why the committee made the recommendations: Procedures that were not recommended

Transthoracic needle biopsy, bronchoscopy and non-ultrasound-guided TBNA are no longer recommended for diagnosing and staging lung cancer in mediastinal lymph nodes because:

- bronchoscopy and non-ultrasound-guided TBNA are unlikely to reach the minimum sensitivity required by the British Thoracic Quality Standards **and**
- they may discourage people from having more effective procedures (such as EBUS-TBNA) and subsequent investigations.

The word ‘fibreoptic’ has been removed because bronchoscopy can be fibreoptic, video or hybrid.

Impact of the recommendations on practice

The recommendations on PET-CT reflect current practice, so will not incur an extra cost.

EBUS-TBNA and EUS-FNA are widely used. The recommendations will reinforce best practice and result in a more streamlined diagnostic service with more timely diagnosis.

The surgical mediastinal staging recommendation will also reinforce best practice and restrict this procedure to people most likely to benefit.

Interpreting the evidence

The outcomes that matter most

The committee highlighted that the outcomes that matter most are time to treatment decision, number of investigations per patient, patient acceptability, reduction of avoidable thoracic surgery and diagnostic sensitivity and negative predictive value This is because the committee agreed that these two diagnostic accuracy measurements are the ones that that matter most to clinicians and people with suspected / confirmed lung cancer.

The committee agreed that the outcomes in Kang 2014 (adverse events, patient satisfaction, sensitivity and negative predictive value) are less relevant because both arms of the trial involve giving patients 3 endoscopic interventions. This is less relevant because in the UK, healthcare professionals aim to use fewer endoscopic interventions.

The quality of the evidence

The committee agreed that the quality of evidence for using EBUS-TBNA as a first invasive test was good particularly with regard to the study by Navani et al. (2015). The committee also confirmed that the evidence for when EUS-FNA should be used as a first invasive test or as a second invasive test following EBUS-TBNA was of a lower quality: The methods section of Navani 2015 says the following: "If a target node was inaccessible with EBUS-TBNA then EUS-FNA as an alternative procedure was allowed." The word "inaccessible" is an inexact term. For example, this term does not specify which lung stations are inaccessible by EBUS-TBNA. In Navani

2015, EUS-FNA was conducted for 2 people who met the inclusion criteria out of 66 (the others had EBUS-TBNA because they had suspicious lesions in lung stations accessible by EBUS-TBNA). To specify a more exact treatment protocol that includes EUS-FNA, there is an issue of collecting enough data. Therefore, the committee agreed that it might never be possible to have a study that specifies the exact usage of EUS-FNA. This is because the outcomes depend on too many variables such as the study population. In addition, Kang 2014 had vague inclusion criteria, non-significant results and had indirect evidence because the in the UK clinicians aim to give patients fewer than 3 endoscopic interventions. The committee also noted that EUS-FNA is particularly good at reaching lung stations 8, 9 and 4L.

Benefits and harms

The committee agreed that EBUS-TBNA and/or EUS-FNA should be offered as a first invasive test for diagnosis and staging lung cancer with a probability of having mediastinal malignancy. This is because the committee decided that the findings of Navani 2015 showed that for EBUS-TBNA (or EUS-FNA) there was a reduction in time to treatment decision, a reduction in the number of investigations per patient and an increase in the number of people diagnosed and staged with one investigation compared to conventional diagnosis and staging (bronchoscopy or CT-guided biopsy etc.). The committee also found it plausible that the higher rates of survival in the EBUS-TBNA arm of the trial might be related to the faster treatment decisions those patients received. In addition, the committee noted that the findings in Annema 2010 and Larsen 2005 show that EBUS-TBNA and/or EUS-FNA as a first invasive test for people with a probability of having mediastinal malignancy, reduces the number of avoidable thoracic surgeries compared to people who go straight to surgical staging. Finally, EBUS-TBNA and EUS-FNA are generally performed as day case procedures under sedation and are safer, faster, cheaper and repeatable if necessary compared to surgical staging. The committee decided to recommend that EBUS-TBNA and EUS-FNA be offered together where indicated as this would be better for patients and consume less resources than if the two procedures were performed on separate occasions.

Cost effectiveness and resource use

The committee examined cost data on the various procedures and acknowledged that although it was recognised to be less sensitive than EBUS-TBNA, conventional TBNA would be the cheaper option for accessing lymph nodes via the trachea. They noted, however, that the large apparent cost differences between conventional TBNA and EBUS-TBNA are an artefact of certain pricing codes used in published sources (Luque et al. 2016 and the 2011 version of this guideline) and likely to be far smaller in reality, as the only difference between the procedures are the marginal costs associated with the EBUS equipment and the difference between the costs of the needles. This was calculated at a little over £300 per procedure (see Appendix J). In addition many NHS trusts already have the EBUS equipment

The committee considered whether they should recommend a cost saving strategy that put conventional TBNA first in a sequenced diagnostic pathway, followed by EBUS-TBNA for patients testing negative. The committee rejected this for several reasons. Firstly, the committee recognised the direction of travel in NHS policy is for time-to-diagnosis to be significantly reduced (a 28 day wait is to be trialled between 2018 and is intended to become national policy by 2020). Secondly, they noted that the National Optimal Lung Cancer Pathway recently published by the Lung Clinical Expert Group recommends that biopsy results should be available to the MDT within 21 days of initial suspicion of lung cancer on a CT scan. Thirdly, they recognised the practical difficulty of scheduling multiple tests for patients within this short time window and also took into account the views of patient representatives, which highlighted the importance of reducing the distressing wait for a diagnosis. Fourthly, the committee took into account patient representatives' unease about undergoing multiple uncomfortable tests, which often require recovery time in a hospital bed. As noted above, the committee had experience of some patients being reluctant to return for further tests if the initial test was negative. Also as noted above, the committee found it plausible that extending time to diagnosis, even by a short time, may adversely affect treatment outcomes.

The committee also considered the cost-utility analysis in the Sharples et al. 2012 study and agreed that due to similar QALY estimates for EBUS/EUS and surgical staging, the analysis would reduce to a cost-comparison as concluded by the paper authors. However, they did not have confidence in the costing of endosonography in the Sharples et al. 2012 study as presented because the combined cost of EBUS-TBNA and EUS-FNA was less than the cost of EBUS-TBNA alone that had been provided in other sources produced at a similar time (the NICE 2011 Lung Cancer guideline update and in Navani et al. 2012).The committee also considered the influence diagram model by Luque et al. (2016), which suggested using cheaper tests before EBUS-TBNA but disregarded the evidence due to lack of face validity in the model's diagnostic accuracy and cost data, particularly for conventional TBNA, which was costed at €80 rather than the ~£1,200 estimated for this update (see Appendix J).

Other factors the committee took into account

The committee gave special consideration to people living in deprived areas. This is because socioeconomic status was identified as a potential equality issue in the equity impact assessment. However, the committee agreed that no additional recommendations were necessary. The committee did not have any reason to believe that the interventions work better or worse in different groups. In addition, there was no data available specific to this population.

29 # Appendix A – Review protocols

30 **Review protocol for the clinical and cost effectiveness of using non-ultrasound-guided TBNA, EBUS-TBNA or EUS-FNA as**
31 **the first invasive test for people with a probability of mediastinal malignancy**

32

Field (based on PRISMA-P)	Content
Review question	What is the clinical and cost effectiveness of using non-ultrasound-guided TBNA, EBUS-TBNA or EUS-FNA as the first invasive test for people with a probability of mediastinal malignancy?
Type of review question	Diagnostic and intervention
Objective of the review	This area was identified as requiring updating during the 2016 surveillance review. It is anticipated that recommendation on the use of non-ultrasound-guided TBNA, EBUS-TBNA or EUS-FNA will be affected.
Eligibility criteria – population	Patients with suspected/ confirmed lung cancer (Pre-diagnosis and CT std. clinical evaluation) or in other words, people with a probability of mediastinal malignancy

Eligibility criteria – interventions	• Non-ultrasound-guided TBNA, • EBUS-TBNA or • EUS-FNA
Eligibility criteria – gold standard	The gold standard investigation (histological/ cytological confirmation and pathological TNM - Or follow up period adequate to confirm outcome - Normally pathology from surgical resection but could be another technique in specified circumstances.
Outcomes and prioritisation	• The diagnostic sensitivity and specificity (likelihood ratios) • The staging sensitivity and specificity • The safety of each procedure/ adverse events (EBUS – mortality, in-patient admission, pneumothorax) • Patient acceptability • Anxiety and psychological outcomes • Timing (e.g. time to treatment) • The number of investigations and outpatient attendances per patient
Eligibility criteria – study design	• RCTs • Systematic review of RCTs • If insufficient evidence is identified, diagnostic cross-sectional studies will be considered.

Other inclusion exclusion criteria	• Non- English-language papers • Unpublished evidence/ conference proceedings
Proposed sensitivity/sub-group analysis, or meta-regression	No subgroup analysis identified
Selection process – duplicate screening/selection/analysis	10% of the abstracts were reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. If meaningful disagreements were found between the different reviewers, a further 10% of the abstracts were reviewed by two reviewers, with this process continued until agreement is achieved between the two reviewers. From this point, the remaining abstracts will be screened by a single reviewer. This review made use of the priority screening functionality with the EPPI-reviewer systematic reviewing software. See Appendix B for more details.
Data management (software)	See Methods Appendix B
Information sources – databases and dates	See Appendix C Main Searches: • Cochrane Database of Systematic Reviews – CDSR • Cochrane Central Register of Controlled Trials – CENTRAL • Database of Abstracts of Reviews of Effects – DARE • Health Technology Assessment Database – HTA • EMBASE (Ovid) • MEDLINE (Ovid)

	• MEDLINE In-Process (Ovid) Citation searching will be carried out in addition on analyst/committee selected papers. The search will not be date limited because this is a new review question. Economics: • NHS Economic Evaluation Database – NHS EED • Health Economic Evaluations Database – HEED • EconLit (Ovid) • Embase (Ovid) • MEDLINE (Ovid) • MEDLINE In-Process (Ovid) The search will not be date limited because this is a new review question.
Identify if an update	This is not an update, this is a new review question.
Author contacts	Guideline update
Highlight if amendment to previous protocol	For details please see section 4.5 of Developing NICE guidelines: the manual
Search strategy – for one database	For details please see appendix C

Data collection process – forms/duplicate	A standardised evidence table format will be used, and published as appendix F (clinical evidence tables) or I (economic evidence tables).
Data items – define all variables to be collected	For details please see evidence tables in appendix F (clinical evidence tables) or I (economic evidence tables).
Methods for assessing bias at outcome/study level	See Appendix B
Criteria for quantitative synthesis	See Appendix B
Methods for quantitative analysis – combining studies and exploring (in)consistency	See Appendix B
Meta-bias assessment – publication bias, selective reporting bias	See Appendix B
Confidence in cumulative evidence	See Appendix B
Rationale/context – what is known	For details please see the introduction to the evidence review in the main file.
Describe contributions of authors and guarantor	A multidisciplinary committee developed the evidence review. The committee was convened by the NICE Guideline Updates Team and chaired by Gary McVeigh in line with section 3 of Developing NICE guidelines: the manual. Staff from the NICE Guideline Updates Team undertook systematic literature searches, appraised the evidence, conducted meta-

	analysis and cost-effectiveness analysis where appropriate, and drafted the evidence review in collaboration with the committee. For details please see Developing NICE guidelines: the manual.
Sources of funding/support	The NICE Guideline Updates Team is an internal team within NICE.
Name of sponsor	The NICE Guideline Updates Team is an internal team within NICE.
Roles of sponsor	The NICE Guideline Updates Team is an internal team within NICE.
PROSPERO registration number	N/A

Review protocol for the clinical and cost-effectiveness of EBUS-TBNA alone, EUS-FNA alone or EBUS-TBNA and EUS-FNA in combination compared with surgical staging to diagnose and/or stage lung cancer

What is the clinical and cost-effectiveness of EBUS-TBNA alone, EUS-FNA alone or EBUS-TBNA and EUS-FNA in combination compared with surgical staging to diagnose and/or stage lung cancer?

Field (based on PRISMA-P)	Content
Review question	What is the clinical and cost-effectiveness of EBUS-TBNA alone, EUS-FNA alone or EBUS-TBNA and EUS-FNA in combination

	compared with surgical staging to diagnose and/or stage lung cancer?
Type of review question	Diagnostic and intervention
Objective of the review	This area was identified as requiring updating during the 2016 surveillance review. Anticipated recommendations may cover which test is most appropriate for diagnosing or staging of lung cancer.
Eligibility criteria – population	Patients with suspected/ confirmed lung cancer (Pre-diagnosis and CT std. clinical evaluation)
Eligibility criteria – interventions	• EBUS-TBNA alone, • EUS-FNA alone or • EBUS-TBNA and EUS-FNA in combination
Eligibility criteria – gold standard	• Surgical staging Or follow up period adequate to confirm outcome - Normally pathology from surgical resection but could be another technique in specified circumstances.

Outcomes and prioritisation	• The diagnostic sensitivity and specificity (likelihood ratios) • The staging sensitivity and specificity • The safety of each procedure/ adverse events (EBUS – mortality, in-patient admission, pneumothorax) • Patient acceptability • Anxiety and psychological outcomes – report if in evidence • Quality of life • The number of investigations and outpatient attendances per patient • Timing (e.g. time to treatment)
Eligibility criteria – study design	• RCTs • Systematic review of RCTs • If insufficient evidence is identified, diagnostic cross-sectional studies will be considered.
Other inclusion exclusion criteria	• Non- English-language papers • Unpublished evidence/ conference proceedings
Proposed sensitivity/sub-group analysis, or meta-regression	No subgroup analysis identified
Selection process – duplicate screening/selection/analysis	10% of the abstracts were reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third

	independent reviewer. If meaningful disagreements were found between the different reviewers, a further 10% of the abstracts were reviewed by two reviewers, with this process continued until agreement is achieved between the two reviewers. From this point, the remaining abstracts will be screened by a single reviewer. This review made use of the priority screening functionality with the EPPI-reviewer systematic reviewing software. See Appendix B for more details.
Data management (software)	See Methods Appendix B
Information sources – databases and dates	See Appendix C Main Searches: • Cochrane Database of Systematic Reviews – CDSR • Cochrane Central Register of Controlled Trials – CENTRAL • Database of Abstracts of Reviews of Effects – DARE • Health Technology Assessment Database – HTA

	• EMBASE (Ovid) • MEDLINE (Ovid) • MEDLINE In-Process (Ovid) Citation searching will be carried out in addition on analyst/committee selected papers. The search will not be date limited because this is a new review question. Economics: • NHS Economic Evaluation Database – NHS EED • Health Economic Evaluations Database – HEED • EconLit (Ovid) • Embase (Ovid) • MEDLINE (Ovid) • MEDLINE In-Process (Ovid) The search will not be date limited because this is a new review question.
Identify if an update	This is not an update, this is a new review question.
Author contacts	Guideline update

Highlight if amendment to previous protocol	For details please see section 4.5 of [Developing NICE guidelines: the manual](#)
Search strategy – for one database	For details please see appendix C
Data collection process – forms/duplicate	A standardised evidence table format will be used, and published as appendix E (clinical evidence tables) or I (economic evidence tables).
Data items – define all variables to be collected	For details please see evidence tables in appendix E (clinical evidence tables) or I (economic evidence tables).
Methods for assessing bias at outcome/study level	See Appendix B
Criteria for quantitative synthesis	See Appendix B
Methods for quantitative analysis – combining studies and exploring (in)consistency	See Appendix B
Meta-bias assessment – publication bias, selective reporting bias	See Appendix B
Confidence in cumulative evidence	See Appendix B
Rationale/context – what is known	For details please see the introduction to the evidence review in the main file.

Describe contributions of authors and guarantor	A multidisciplinary committee developed the evidence review. The committee was convened by the NICE Guideline Updates Team and chaired by Gary McVeigh in line with section 3 of Developing NICE guidelines: the manual. Staff from the NICE Guideline Updates Team undertook systematic literature searches, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate, and drafted the evidence review in collaboration with the committee. For details please see Developing NICE guidelines: the manual.
Sources of funding/support	The NICE Guideline Updates Team is an internal team within NICE.
Name of sponsor	The NICE Guideline Updates Team is an internal team within NICE.
Roles of sponsor	The NICE Guideline Updates Team is an internal team within NICE.

PROSPERO registration number	N/A

Appendix B – Methods

Priority screening

The reviews undertaken for this guideline all made use of the priority screening functionality with the EPPI-reviewer systematic reviewing software. This uses a machine learning algorithm (specifically, an SGD classifier) to take information on features (1, 2 and 3 word blocks) in the titles and abstract of papers marked as being ‘includes’ or ‘excludes’ during the title and abstract screening process, and re-orders the remaining records from most likely to least likely to be an include, based on that algorithm. This re-ordering of the remaining records occurs every time 25 additional records have been screened.

Research is currently ongoing as to what are the appropriate thresholds where reviewing of abstract can be stopped, assuming a defined threshold for the proportion of relevant papers it is acceptable to miss on primary screening. As a conservative approach until that research has been completed, the following rules were adopted during the production of this guideline:

- In every review, at least 50% of the identified abstract (or 1,000 records, if that is a greater number) were always screened.
- After this point, screening was only terminated when the threshold was reached for a number of abstracts being screened without a single new include being identified. This threshold was set according to the expected proportion of includes in the review (with reviews with a lower proportion of includes needing a higher number of papers without an identified study to justify termination), and was always a minimum of 250.
- A random 10% sample of the studies remaining in the database when the threshold were additionally screened, to check if a substantial number of relevant studies were not being correctly classified by the algorithm, with the full database being screened if concerns were identified.

As an additional check to ensure this approach did not miss relevant studies, the included studies lists of included systematic reviews were searched to identify any papers not identified through the primary search.

Evidence synthesis and meta-analyses

Where possible, meta-analyses were conducted to combine the results of studies for each outcome. For mean differences, where change from baseline data were reported in the studies and were accompanied by a measure of spread (for example standard deviation), these were extracted and used in the meta-analysis. Where measures of spread for change from baseline values were not reported, the corresponding values at study end were used and were combined with change from baseline values to produce summary estimates of effect. All studies were assessed to ensure that baseline values were balanced across the treatment/comparison groups; if there were significant differences in important confounding variables at baseline these studies were not included in any meta-analysis and were reported separately.

When averages were given as medians, no meta-analysis of the data were performed.

Evidence of effectiveness of interventions

Quality assessment

Individual RCTs and quasi-randomised controlled trials were quality assessed using the Cochrane Risk of Bias Tool. Cohort studies were quality assessed using the CASP cohort study checklist. Each individual study was classified into one of the following three groups:

- Low risk of bias – The true effect size for the study is likely to be close to the estimated effect size.
- Moderate risk of bias – There is a possibility the true effect size for the study is substantially different to the estimated effect size.
- High risk of bias – It is likely the true effect size for the study is substantially different to the estimated effect size.

Each individual study was also classified into one of three groups for directness, based on if there were concerns about the population, intervention, comparator and/or outcomes in the study and how directly these variables could address the specified review question. Studies were rated as follows:

- Direct – No important deviations from the protocol in population, intervention, comparator and/or outcomes.
- Partially indirect – Important deviations from the protocol in one of the population, intervention, comparator and/or outcomes.
- Indirect – Important deviations from the protocol in at least two of the following areas: population, intervention, comparator and/or outcomes.

Methods for combining intervention evidence

Meta-analyses of interventional data were conducted with reference to the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al. 2011).

Where different studies presented continuous data measuring the same outcome but using different numerical scales (e.g. a 0-10 and a 0-100 visual analogue scale), these outcomes were all converted to the same scale before meta-analysis was conducted on the mean differences. Where outcomes measured the same underlying construct but used different instruments/metrics, data were analysed using standardised mean differences (Hedges' g).

A pooled relative risk was calculated for dichotomous outcomes (using the Mantel–Haenszel method). Both relative and absolute risks were presented, with absolute risks calculated by applying the relative risk to the pooled risk in the comparator arm of the meta-analysis.

Fixed- and random-effects models (der Simonian and Laird) were fitted for all syntheses, with the presented analysis dependent on the degree of heterogeneity in the assembled evidence. Fixed-effects models were the preferred choice to report, but in situations where the assumption of a shared mean for fixed-effects model were clearly not met, even after appropriate pre-specified subgroup analyses were conducted, random-effects results are presented. Fixed-effects models were deemed to be inappropriate if one or both of the following conditions was met:

- Significant between study heterogeneity in methodology, population, intervention or comparator was identified by the reviewer in advance of data analysis. This decision was made and recorded before any data analysis was undertaken.

- The presence of significant statistical heterogeneity in the meta-analysis, defined as $I^2 \geq 50\%$.

In any meta-analyses where some (but not all) of the data came from studies at high risk of bias, a sensitivity analysis was conducted, excluding those studies from the analysis. Results from both the full and restricted meta-analyses are reported. Similarly, in any meta-analyses where some (but not all) of the data came from indirect studies, a sensitivity analysis was conducted, excluding those studies from the analysis.

Meta-analyses were performed in Cochrane Review Manager v 5.3.

Minimal clinically important differences (MIDs)

The Core Outcome Measures in Effectiveness Trials (COMET) database was searched to identify published minimal clinically important difference thresholds relevant to this guideline. However, no relevant MIDs were found. In addition, the Guideline Committee were asked to specify any outcomes where they felt a consensus MID could be defined from their experience. In particular, any questions looking to evaluate non-inferiority (that one intervention is not meaningfully worse than another) required an MID to be defined to act as a non-inferiority margin. However, the committee agreed that in their experience, they could not define any MIDs. This is because the committee agreed that the protocol outcomes were objective rather than subjective measures and the committee were not aware of evidence supporting the use of MIDs for the protocol's outcomes. This was particularly the case for sensitivity and negative predictive value. The line of no effect was used as a MID for risk ratios and hazard ratios. Diagnostic accuracy outcomes do not have a line of no effect. Therefore, imprecision for diagnostic accuracy was graded using participant numbers only.

GRADE for pairwise meta-analyses of interventional evidence

GRADE was used to assess the quality of evidence for the selected outcomes as specified in 'Developing NICE guidelines: the manual (2014)'. Data from RCTs was initially rated as high quality and the quality of the evidence for each outcome was downgraded or not from this initial point. If non-RCT evidence was included for intervention-type systematic reviews then these were initially rated as either moderate quality (quasi-randomised studies) or low quality (cohort studies) and the quality of the evidence for each outcome was further downgraded or not from this point, based on the criteria given in Table 4.

Table 4: Rationale for downgrading quality of evidence for intervention studies

GRADE criteria	Reasons for downgrading quality
Risk of bias	Not serious: If less than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the overall outcome was not downgraded. Serious: If greater than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the outcome was downgraded one level. Very serious: If greater than 33.3% of the weight in a meta-analysis came from studies at high risk of bias, the outcome was downgraded two levels. Outcomes meeting the criteria for downgrading above were not downgraded if there was evidence the effect size was not meaningfully different between studies at high and low risk of bias.
Indirectness	Not serious: If less than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the overall outcome was not downgraded.

GRADE criteria	Reasons for downgrading quality
	Serious: If greater than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the outcome was downgraded one level. Very serious: If greater than 33.3% of the weight in a meta-analysis came from indirect studies, the outcome was downgraded two levels. Outcomes meeting the criteria for downgrading above were not downgraded if there was evidence the effect size was not meaningfully different between direct and indirect studies.
Inconsistency	Concerns about inconsistency of effects across studies, occurring when there is unexplained variability in the treatment effect demonstrated across studies (heterogeneity), after appropriate pre-specified subgroup analyses have been conducted. This was assessed using the I^2 statistic. N/A: Inconsistency was marked as not applicable if data on the outcome was only available from one study. Not serious: If the I^2 was less than 33.3%, the outcome was not downgraded. Serious: If the I^2 was between 33.3% and 66.7%, the outcome was downgraded one level. Very serious: If the I^2 was greater than 66.7%, the outcome was downgraded two levels. Outcomes meeting the criteria for downgrading above were not downgraded if there was evidence the effect size was not meaningfully different between studies with the smallest and largest effect sizes.
Imprecision	The line of no effect was defined as the MID for risk ratios and hazard ratios. Risk ratios and hazard ratios were downgraded once if the 95% confidence interval of the effect size crossed the line of no effect. For pooled mean differences, a MID of 0.2 SD was used. If the 95% confidence interval of the effect size crossed one line of no effect, the outcome was downgraded once. If the 95% confidence interval crossed both lines of no effect, the outcome was downgraded twice. The committee agreed that if the sample size was 26 to 40, the outcome was downgraded once. If the sample size was 25 or less, the outcome was downgraded twice. Outcomes meeting the criteria for downgrading above were not downgraded if the confidence interval was sufficiently narrow that the upper and lower bounds would correspond to clinically equivalent scenarios.

The quality of evidence for each outcome was upgraded if any of the following five conditions were met:

- Data from non-randomised studies showing an effect size sufficiently large that it cannot be explained by confounding alone.
- Data showing a dose-response gradient.
- Data where all plausible residual confounding is likely to increase our confidence in the effect estimate.

Publication bias

Publication bias was assessed in two ways. First, if evidence of conducted but unpublished studies was identified during the review (e.g. conference abstracts, trial protocols or trial records without accompanying published data), available information on these unpublished studies was reported as part of the review. Secondly, where 10 or more studies were included as part of a single meta-analysis, a funnel plot was produced to graphically assess the potential for publication bias.

Evidence statements

Evidence statements for pairwise intervention data are classified in to one of four categories:

- Situations where the data are only consistent, at a 95% confidence level, with an effect in one direction (i.e. one that is 'statistically significant'), and the magnitude of that effect is most likely to meet or exceed the MID (i.e. the point estimate is not in the zone of equivalence). In such cases, we state that the evidence showed that there is an effect.
- Situations where the data are only consistent, at a 95% confidence level, with an effect in one direction (i.e. one that is 'statistically significant'), but the magnitude of that effect is most likely to be less than the MID (i.e. the point estimate is in the zone of equivalence). In such cases, we state that the evidence could not demonstrate a meaningful difference.
- Situations where the data are consistent, at a 95% confidence level, with an effect in either direction (i.e. one that is not 'statistically significant') but the confidence limits are smaller than the MIDs in both directions. In such cases, we state that the evidence demonstrates that there is no difference.
- In all other cases, we state that the evidence could not differentiate between the comparators.

Diagnostic test accuracy evidence

In this guideline, diagnostic test accuracy (DTA) data are classified as any data in which a test result or the output of an algorithm – is observed in some people who have the condition of interest at the time of the test and some people who do not. Such data either explicitly provide, or can be manipulated to generate, a 2x2 classification of true positives and false negatives (in people who, according to the reference standard, truly have the condition) and false positives and true negatives (in people who, according to the reference standard, do not).

The 'raw' 2x2 data can be summarised in a variety of ways. Those that were used for decision making in this guideline are as follows:

- **Sensitivity** is the probability that the feature will be positive in a person with the condition.
  - sensitivity = TP/(TP+FN)
- **Negative predictive value** is the probability that people for whom the feature is negative truly do not have the condition.
  - negative predictive value = TN/(TN+FN)

Negative predictive value was used rather than specificity. This is because all studies assumed that the pathologist made no false positives. Therefore, sensitivity and negative predictive value (with prevalence information) are more meaningful measurements of performance because they do not involve false positives.

Quality assessment

Individual studies were quality assessed using the QUADAS-2 tool, which contains four domains: patient selection, index test, reference standard, and flow and timing. Each individual study was classified into one of the following two groups:

- Low risk of bias – Evidence of non-serious bias in zero or one domain.
- Moderate risk of bias – Evidence of non-serious bias in two domains only, or serious bias in one domain only.

- High risk of bias – Evidence of bias in at least three domains, or of serious bias in at least two domains.

Each individual study was also classified into one of three groups for directness, based on if there were concerns about the population, index features and/or reference standard in the study and how directly these variables could address the specified review question. Studies were rated as follows:

- Direct – No important deviations from the protocol in population, index feature and/or reference standard.
- Partially indirect – Important deviations from the protocol in one of the population, index feature and/or reference standard.
- Indirect – Important deviations from the protocol in at least two of the population, index feature and/or reference standard.

Modified GRADE for diagnostic test accuracy evidence

GRADE has not been developed for use with diagnostic studies; therefore a modified approach was applied using the GRADE framework. GRADE assessments were only undertaken for sensitivity and negative predictive values (that are provided in the context of the prevalences of lung cancer). The committee thought that it was very unlikely that pathologists would identify non-cancerous cells as cancerous. Therefore, the committee agreed that the false positive rate for all techniques was likely to be 0. Therefore, all calculated outcomes that involve a false positive value are not meaningful. For example, specificity and likelihood ratios. GRADE quality ratings were calculated using the same criteria as for randomised controlled trials, given in Table 4. For example, the committee agreed that if the sample size was 26 to 40, the outcome was downgraded once. If the sample size was 25 or less, the outcome was downgraded twice. This is because neither sensitivity nor negative predictive value have a line of no effect with which to rate imprecision.

Appendix C – Literature search strategies

Scoping search strategies

Scoping searches Scoping searches were undertaken on the following websites and databases (listed in alphabetical order) in April 2017 to provide information for scope development and project planning. Browsing or simple search strategies were employed.

Guidelines/website
American Cancer Society
American College of Chest Physicians
American Society for Radiation Oncology
American Thoracic Society
Association for Molecular Pathology
British Lung Foundation
British Thoracic Society
Canadian Medical Association Infobase
Canadian Task Force on Preventive Health Care
Cancer Australia
Cancer Care Ontario
Cancer Control Alberta
Cancer Research UK
Care Quality Commission
College of American Pathologists
Core Outcome Measures in Effectiveness Trials (COMET)
Department of Health & Social Care
European Respiratory Society
European Society for Medical Oncology
European Society of Gastrointestinal Endoscopy
European Society of Thoracic Surgery
General Medical Council
Guidelines & Audit Implementation Network (GAIN)
Guidelines International Network (GIN)
Healthtalk Online
International Association for the Study of Lung Cancer
MacMillan Cancer Support
Medicines and Products Regulatory Agency (MHRA)
National Audit Office
National Cancer Intelligence Network
National Clinical Audit and Patient Outcomes Programme
National Health and Medical Research Council - Australia
National Institute for Health and Care Excellence (NICE) - published & in development guidelines
National Institute for Health and Care Excellence (NICE) - Topic Selection
NHS Choices
NHS Digital
NHS England
NICE Clinical Knowledge Summaries (CKS)

Guidelines/website
NICE Evidence Search
Office for National Statistics
Patient UK
PatientVoices
Public Health England
Quality Health
Royal College of Anaesthetists
Royal College of General Practitioners
Royal College of Midwives
Royal College of Nursing
Royal College of Pathologists
Royal College of Physicians
Royal College of Radiologists
Royal College of Surgeons
Scottish Government
Scottish Intercollegiate Guidelines Network (SIGN)
UK Data Service
US National Guideline Clearinghouse
Walsall community Health NHS Trust
Welsh Government

Clinical search literature search strategy

Main searches

Bibliographic databases searched for the guideline

- Cochrane Database of Systematic Reviews – CDSR (Wiley)
- Cochrane Central Register of Controlled Trials – CENTRAL (Wiley)
- Database of Abstracts of Reviews of Effects – DARE (Wiley)
- Health Technology Assessment Database – HTA (Wiley)
- EMBASE (Ovid)
- MEDLINE (Ovid)
- MEDLINE Epub Ahead of Print (Ovid)
- MEDLINE In-Process (Ovid)

Identification of evidence for review questions

The searches were conducted between October 2017 and April 2018 for 9 review questions (RQ).

Searches were re-run in May 2018.

Where appropriate, in-house study design filters were used to limit the retrieval to, for example, randomised controlled trials. Details of the study design filters used can be found in section 3.

Search strategy

Medline Strategy, searched 3rd November 2017 Database: Ovid MEDLINE(R) 1946 to October Week 4 2017 Search Strategy:
1 exp Lung Neoplasms/
2 ((lung* or pulmonary or bronch*) adj3 (cancer* or neoplasm* or carcinoma* or tumo?r* or lymphoma* or metast* or malignan* or blastoma* or carcinogen* or adenocarcinoma* or angiosarcoma* or chrondosarcoma* or sarcoma* or teratoma* or microcytic*)).tw.
3 ((pancoast* or superior sulcus or pulmonary sulcus) adj4 (tumo?r* or syndrome*)).tw.
4 ((lung* or pulmonary or bronch*) adj4 (oat or small or non-small) adj4 cell*).tw.
5 (SCLC or NSCLC).tw.
6 or/1-5
7 exp Biopsy, Fine-Needle/
8 Biopsy, Needle/mt [Methods]
9 (TBNA* or EBUSTBNA* or TBNB* or EUS-FNA* or EUSFNA* or EUS-FNB* or EUSFNB*).tw.
10 (EUS* adj2 (FNA* or FNB*)).tw.
11 ((transbronch* or trans-bronch*) adj4 needle* adj4 (aspirat* or biops* or prick* or perforat* or ruptur*)).tw.
12 ((endoscop* or endobronch*) adj4 (ultras* or echo* or sonogra* or tomograph* or doptone*) adj4 (needle* or fine or hollow*) adj4 (aspirat* or biops* or prick* or perforat* or ruptur*)).tw.
13 (EUS* adj4 (needle* or fine or hollow*) adj4 (aspirat* or biops* or prick* or perforat* or ruptur*)).tw.
14 or/7-13
15 6 and 14
16 Animals/ not Humans/
17 15 not 16
18 limit 17 to english language

Note: *In-house RCT, observational studies and systematic review filters were appended. No date limit as these were new questions.*

Study Design Filters

The MEDLINE SR, RCT, and observational studies filters are presented below.
Systematic Review
1. Meta-Analysis.pt.
2. Meta-Analysis as Topic/
3. Review.pt.
4. exp Review Literature as Topic/
5. (metaanaly$ or metanaly$ or (meta adj3 analy$)).tw.
6. (review$ or overview$).ti.
7. (systematic$ adj5 (review$ or overview$)).tw.
8. ((quantitative$ or qualitative$) adj5 (review$ or overview$)).tw.
9. ((studies or trial$) adj2 (review$ or overview$)).tw.
10. (integrat$ adj3 (research or review$ or literature)).tw.
11. (pool$ adj2 (analy$ or data)).tw.
12. (handsearch$ or (hand adj3 search$)).tw.
13. (manual$ adj3 search$).tw.

The MEDLINE SR, RCT, and observational studies filters are presented below.
14. or/1-13 15. animals/ not humans/ 16. 14 not 15
RCT 1 Randomized Controlled Trial.pt. 2 Controlled Clinical Trial.pt. 3 Clinical Trial.pt. 4 exp Clinical Trials as Topic/ 5 Placebos/ 6 Random Allocation/ 7 Double-Blind Method/ 8 Single-Blind Method/ 9 Cross-Over Studies/ 10 ((random$ or control$ or clinical$) adj3 (trial$ or stud$)).tw. 11 (random$ adj3 allocat$).tw. 12 placebo$.tw. 13 ((singl$ or doubl$ or trebl$ or tripl$) adj (blind$ or mask$)).tw. 14 (crossover$ or (cross adj over$)).tw. 15 or/1-14 16 animals/ not humans/ 17 15 not 16
Observational 1 Observational Studies as Topic/ 2 Observational Study/ 3 Epidemiologic Studies/ 4 exp Case-Control Studies/ 5 exp Cohort Studies/ 6 Cross-Sectional Studies/ 7 Controlled Before-After Studies/ 8 Historically Controlled Study/ 9 Interrupted Time Series Analysis/ 10 Comparative Study.pt. 11 case control$.tw. 12 case series.tw. 13 (cohort adj (study or studies)).tw. 14 cohort analy$.tw. 15 (follow up adj (study or studies)).tw. 16 (observational adj (study or studies)).tw. 17 longitudinal.tw. 18 prospective.tw. 19 retrospective.tw. 20 cross sectional.tw. 21 or/1-20

Health Economics literature search strategy

Sources searched to identify economic evaluations

- NHS Economic Evaluation Database – NHS EED (Wiley) last updated Apr 2015
- Health Technology Assessment Database – HTA (Wiley) last updated Oct 2016

- Embase (Ovid)
- MEDLINE (Ovid)
- MEDLINE In-Process (Ovid)

Search filters to retrieve economic evaluations and quality of life papers were appended to the review question search strategies. For some health economics strategies additional terms were added to the original review question search strategies (see sections 4.2, 4.3 and 4.4) The searches were conducted between October 2017 and April 2018 for 9 review questions (RQ).

Searches were re-run in May 2018.

Searches were limited to those in the English language. Animal studies were removed from results.

Economic evaluation and quality of life filters

Medline Strategy	
Economic evaluations	
1	Economics/
2	exp "Costs and Cost Analysis"/
3	Economics, Dental/
4	exp Economics, Hospital/
5	exp Economics, Medical/
6	Economics, Nursing/
7	Economics, Pharmaceutical/
8	Budgets/
9	exp Models, Economic/
10	Markov Chains/
11	Monte Carlo Method/
12	Decision Trees/
13	econom$.tw.
14	cba.tw.
15	cea.tw.
16	cua.tw.
17	markov$.tw.
18	(monte adj carlo).tw.
19	(decision adj3 (tree$ or analys$)).tw.
20	(cost or costs or costing$ or costly or costed).tw.
21	(price$ or pricing$).tw.
22	budget$.tw.
23	expenditure$.tw.
24	(value adj3 (money or monetary)).tw.
25	(pharmacoeconomic$ or (pharmaco adj economic$)).tw.
26	or/1-25
Quality of life	
1	"Quality of Life"/
2	quality of life.tw.

Medline Strategy
3 "Value of Life"/
4 Quality-Adjusted Life Years/
5 quality adjusted life.tw.
6 (qaly$ or qald$ or qale$ or qtime$).tw.
7 disability adjusted life.tw.
8 daly$.tw.
9 Health Status Indicators/
10 (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).tw.
11 (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw.
12 (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).tw.
13 (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).tw.
14 (sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).tw.
15 (euroqol or euro qol or eq5d or eq 5d).tw.
16 (qol or hql or hqol or hrqol).tw.
17 (hye or hyes).tw.
18 health$ year$ equivalent$.tw.
19 utilit$.tw.
20 (hui or hui1 or hui2 or hui3).tw.
21 disutili$.tw.
22 rosser.tw.
23 quality of wellbeing.tw.
24 quality of well-being.tw.
25 qwb.tw.
26 willingness to pay.tw.
27 standard gamble$.tw.
28 time trade off.tw.
29 time tradeoff.tw.
30 tto.tw.
31 or/1-30

278 Health economics search strategy

Medline Strategy, searched 6th November 2017 **Database: Ovid MEDLINE(R) 1946 to October Week 4 2017** **Search Strategy:**
1 exp Lung Neoplasms/
2 ((lung* or pulmonary or bronch*) adj3 (cancer* or neoplasm* or carcinoma* or tumo?r* or lymphoma* or metast* or malignan* or blastoma* or carcinogen* or adenocarcinoma* or angiosarcoma* or chrondosarcoma* or sarcoma* or teratoma* or microcytic*)).tw.
3 ((pancoast* or superior sulcus or pulmonary sulcus) adj4 (tumo?r* or syndrome*)).tw. (756)
4 ((lung* or pulmonary or bronch*) adj4 (oat or small or non-small) adj4 cell*).tw.
5 (SCLC or NSCLC).tw.
6 or/1-5

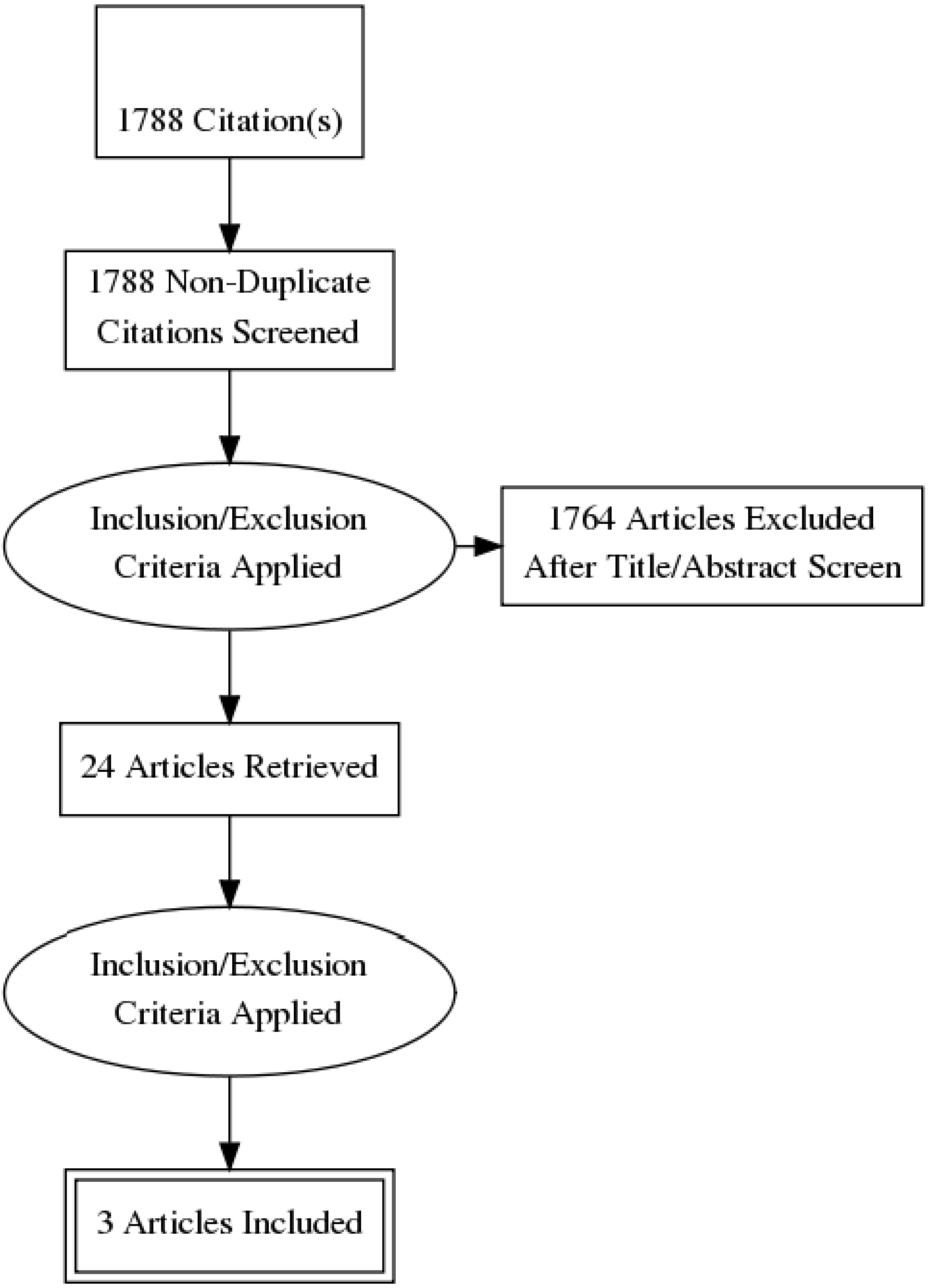
Medline Strategy, searched 6th November 2017
Database: Ovid MEDLINE(R) 1946 to October Week 4 2017
Search Strategy:

7 exp Biopsy, Fine-Needle/
8 Biopsy, Needle/mt [Methods]
9 (TBNA* or EBUSTBNA* or TBNB* or EUS-FNA* or EUSFNA* or EUS-FNB* or EUSFNB*).tw.
10 (EUS* adj2 (FNA* or FNB*)).tw.
11 ((transbronch* or trans-bronch*) adj4 needle* adj4 (aspirat* or biops* or prick* or perforat* or ruptur*)).tw.
12 ((endoscop* or endobronch*) adj4 (ultras* or echo* or sonogra* or tomograph* or doptone*) adj4 (needle* or fine or hollow*) adj4 (aspirat* or biops* or prick* or perforat* or ruptur*)).tw.
13 (EUS* adj4 (needle* or fine or hollow*) adj4 (aspirat* or biops* or prick* or perforat* or ruptur*)).tw.
14 exp Positron-Emission Tomography/
15 (positron emission adj2 compute* adj2 (tomograph* or assist*)).tw.
16 (PET* adj2 CT).tw.
17 Mediastinoscopy/
18 Mediastinoscopes/
19 Mediastinum/dg [Diagnostic Imaging]
20 (mediastinoscop* or mediastinotom*).tw.
21 ((neck* or collum or collar) adj4 US).tw.
22 or/7-21
23 exp Neck/
24 Neck Muscles/
25 exp Cervical Vertebrae/
26 (neck* or collum or collar).tw.
27 ((cervical or C) adj4 vertebra*).tw.
28 or/23-27
29 exp Ultrasonography/
30 (ultras* or echo* or sonogra* or tomograph* or doptone*).tw.
31 29 or 30
32 28 and 31
33 22 or 32
34 6 and 33 (10309)
35 Animals/ not Humans/
36 34 not 35
37 limit 36 to english language

Appendix D – Evidence study selection for RQ 1.1 and RQ 1.2

Clinical evidence study selection

Economic evidence study selection



Appendix E – Clinical evidence tables

Short Title	Title	Study Characteristics	Risk of Bias
Annema 2010	Mediastinoscopy vs endosonography for mediastinal nodal staging of lung cancer: a randomized trial	**Study type** • Randomised controlled trial *This is the ASTER RCT, which has a mirror publication - Sharples 2012. ASTER is short for: Assessment of Surgical sTaging versus Endosonographic ultrasound in lung cancer: a Randomised clinical trial. Data in Sharples 2012 was also used in this analysis.* **Study details** • Study location *Netherlands, Belgium, UK* • Study setting *Leiden University Medical Center, the Netherlands; the University Hospitals of Ghent and Leuven in Belgium; and Papworth Hospital United Kingdom.* • Study dates *February 2007 to April 2009* • Duration of follow-up *Study inclusion, preliminary findings, and complications were evaluated 1 year after start of the study. Patients were followed up for survival for 6 months after staging.* • Sources of funding *Local support for data collection at Ghent University Hospital was provided by the Zorg-programma Oncologie Gent (ZOG) (Ghent University Hospital). Data collection in Papworth Hospital was supported by the UK National Health Service R&D Health. Two of the*	**Quality assessment (RCT)** Random sequence generation • Unclear risk of bias *Details of the randomisation method are not provided.* Allocation concealment • Unclear risk of bias *No mention of allocation concealment.* Blinding of outcome assessment • Unclear risk of bias *No mention of how aware pathologists and radiologists were of the trial taking place.* Blinding of participants and personnel • Unclear risk of bias *Blinding is not possible for a study of this nature.* Incomplete outcome data • Low risk of bias Selective reporting • Low risk of bias

Short Title	Title	Study Characteristics	Risk of Bias
		investigators were supported in part by the National Institute for Health Research Cambridge Biomedical Research Centre. • Lung cancer staging system used *European Society of Thoracic Surgeons Guidelines 2007* **Inclusion criteria** • Suspected N2 or N3 mediastinal lymph node involvement **Exclusion criteria** • <18 years of age • Not fit enough to undergo thoracotomy and lung resection • Significant concurrent malignant disease • Any condition that contraindicated the intervention or mediastinoscopy • Known extrathoracic malignant disease • Received previous treatment for lung cancer • Uncorrected coagulopathy • Unlikely to be staged accurately by any surgical staging procedure • Pregnancy • Inability to consent **Sample characteristics** • Sample size *241 people* • Split between study groups *Straight to surgical staging (mediastinoscopy) = 117 (one person dropped out because they had bone metastasis); EUS-FNA followed by EBUS-TBNA = 123*	Other sources of bias • Low risk of bias Overall risk of bias • Moderate *Details of randomisation are not provided* Directness • Directly applicable **QUADAS 2** Was a random sample of patients enrolled? • Unclear *Details of the randomisation method are not provided.* Was a case-control design avoided? • Yes Did the study avoid inappropriate exclusions? • Yes RISK Could the selection of patients have introduced bias? • Low CONCERN Is there concern that the included patients do not match the review question?

Short Title	Title	Study Characteristics	Risk of Bias
		• Loss to follow-up *All 241 people were followed up.* • %female *Straight to surgical staging = 74% male, 26% female; EUS-FNA then EBUS-TBNA = 80% male, 20% female* • Mean age (SD) *Straight to surgical staging = 65 (9); EUS-FNA then EBUS-TBNA = 65 (9)* • Nodal staging on initial PET/CT scan *Straight to surgical staging = N0: 13%; N1: 14%; N2: 56%; N3: 17%; EUS-FNA then EBUS-TBNA = N0: 7%; N1: 16%; N2: 63%; N3: 13%* **Interventions** • EUS-FNA followed by EBUS-TBNA • Straight to surgical staging (mediastinoscopy) **Downstream investigations and/or treatments** • EUS-FNA followed by EBUS-TBNA arm *58/123 were found to have locally advanced disease. They proceeded to multimodality treatment. 65/123 were without locally advanced disease. They proceeded to surgical staging. 6/65 had locally advanced disease at surgical staging and had multimodality treatment. 59/65 were without locally advanced disease. 58/59 had a thoracotomy. 1/59 had a second endoscopy. Of the 58 who had a thoracotomy, 6/58 had locally advanced disease and 52/58 were without locally advanced disease.* • Straight to surgical staging arm *117/118 went straight to surgical staging. 1/118 did not because they were found to have bone metastasis. At surgical staging, 42/117 had*	• Low Were the index test results interpreted without knowledge of the results of the reference standard? • Unclear *Information about blinding was not provided.* If a threshold was used, was it pre-specified? • Yes RISK Could the conduct or interpretation of the index test have introduced bias? • Unclear Concerns regarding applicability • Low Is the reference standard likely to correctly classify the target condition? • Yes Were the reference standard results interpreted without knowledge of the results of the index test? • Unclear *Details regarding blinding were not provided.* RISK Could the reference standard, its conduct, or its interpretation have introduced bias?

Short Title	Title	Study Characteristics	Risk of Bias
		locally advanced disease. They proceeded to multimodality treatment. 75/117 were without locally advanced disease. Of these, 70/75 underwent thoracotomy, 3/75 refused thoracotomy, 1/75 had endoscopy, 1/75 deteriorated clinically. Of these 75 without locally advanced disease on surgical staging, 16 were found to have locally advanced disease and 59 were found to be without locally advanced disease. **Protocol outcome measures** • Diagnostic sensitivity *Sensitivity = people who the intervention deemed positive [and were confirmed N2/3 by pathology] / (people who the intervention deemed positive [and were confirmed N2/3 by pathology] + people who the intervention deemed negative who were subsequently shown to have N2/3 at thoracotomy [confirmed by pathology])* • Diagnostic negative predictive value *NPV = people who the intervention deemed negative [and were confirmed negative by thoracotomy with pathology] / (people who the intervention deemed negative [and were confirmed negative by thoracotomy with pathology] + people who the intervention deemed negative but had N2/3 as confirmed by thoracotomy and pathology])* • Safety: pneumothorax *This was the only complication that was relevant to EUS-FNA and EBUS-TBNA* • Safety: other complications • Quality of life *The EQ-5D questionnaire was completed using standard proforma at baseline, at the end of staging (after surgical staging but before thoracotomy) and after 2 months and 6 months for all patients recruited at Papworth Hospital. This information was collected for patients in the*	• Unclear CONCERN Is there concern that the target condition as defined by the reference standard does not match the review question? • Low Was there an appropriate interval between index test(s) and reference standard? • Unclear *Timings are not provided.* Did all patients receive a reference standard? • Yes Did patients receive the same reference standard? • Yes Were all patients included in the analysis? • Yes RISK Could the patient flow have introduced bias? • Low Overall quality • Moderate

Short Title	Title	Study Characteristics	Risk of Bias
		continental European centres who were recruited after April 2008. Between February 2007 and April 2008, EQ-5D data were not available from the continental European centres. As this represented a block of time for which no patient completed the EQ-5D, this information was reasonably assumed to be missing at random. **Non-protocol outcome measures** • No. of avoidable thoracotomies *Rate of unnecessary thoracotomies was defined as either exploratory thoracotomy, unexpected presence of mediastinal nodal metastases (pN2/N3) or tumor invasion of the mediastinum at thoracotomy (pT4), pM1, thoracotomy for SCLC or benign disease (other than carcinoid or hamartoma), or death within 30 days after surgery.* • Percentage (or number) of people who died during a specified follow-up period *Patients were followed up for survival for 6 months after staging.*	
Kang 2014	EBUS-centred versus EUS-centred mediastinal staging in lung cancer: a randomised controlled trial	**Study type** • Randomised controlled trial **Study details** • Study location *South Korea* • Study setting *National Cancer Center in Goyang, South Korea* • Study dates *June 2011 to February 2012* • Duration of follow-up *3-5 days after the intervention*	**Quality assessment (RCT)** Random sequence generation • Low risk of bias Allocation concealment • Low risk of bias Blinding of outcome assessment • Unclear risk of bias *Blinding of pathology laboratory staff was not mentioned.* Blinding of participants and personnel

Short Title	Title	Study Characteristics	Risk of Bias
		• Sources of funding *This work was supported by National Cancer Center Grant* • Lung cancer staging system used *Goldstraw P, Crowley J, Chansky K, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. J Thorac Oncol 2007;2:706–14.* **Inclusion criteria** • Histologically confirmed or strongly suspected, potentially operable non-small cell lung cancer **Exclusion criteria** • <18 years of age • Not fit enough to undergo thoracotomy and lung resection • Any condition that contraindicated the intervention or mediastinoscopy • Any medication that contraindicated the intervention or mediastinoscopy • Pregnancy • >80 years of age • M1 disease • Inoperable T4 disease • Mediastinal infiltration or extranodal invasion of the mediastinal lymph node visible on chest CT • Confirmed supraclavicular lymph node metastasis • Pancoast tumours • Ground glass-dominant (>50% in diameter) T1 nodule (≤3 cm)	• Unclear risk of bias *Blinding is not really possible.* Incomplete outcome data • Low risk of bias Selective reporting • Low risk of bias Other sources of bias • Unclear risk of bias *The inclusion criteria are vague with regards to imaging or the standards/guidelines that were used.* Overall risk of bias • Moderate Directness • Indirectly applicable *The inclusion criteria are vague with regards to imaging or guidelines/standards used. In addition, all participants underwent a bronchoscopy just before the interventions of interest.* **QUADAS 2** Was a random sample of patients enrolled? • Yes

Short Title	Title	Study Characteristics	Risk of Bias
		• Drug reaction to lidocaine, midazolam, fentanyl **Sample characteristics** • Sample size *148 people* • Split between study groups *74 in each arm* • Loss to follow-up *None* • %female *Bronchoscopy, then EBUS-TBNA, then – if required – EUS-FNA = 21% female, 79% male; Bronchoscopy, then EUS-FNA, then – if required – EBUS-TBNA = 29% female, 71% male* • Mean age (SD) *Bronchoscopy, then EBUS-TBNA, then – if required – EUS-FNA = 63.21 years (7.91); Bronchoscopy, then EUS-FNA, then – if required – EBUS-TBNA = 62.94 years (8.39)* • Nodal staging on initial PET/CT scan *Bronchoscopy, then EBUS-TBNA, then – if required – EUS-FNA = N0: 35%; N1: 11.25%; N2: 32.5%; N3: 21.25%; Bronchoscopy, then EUS-FNA, then – if required – EBUS-TBNA = N0: 35%; N1: 11.3%; N2: 27.5%; N3: 26.3%* **Interventions** • Bronchoscopy, EBUS-TBNA then EUS-FNA if necessary on mediastinal nodes inaccessible or difficult to access by EBUS-TBNA • Bronchoscopy, EUS-FNA then EBUS-TBNA if necessary on mediastinal nodes inaccessible or difficult to access by EUS-FNA	Was a case-control design avoided? • Yes Did the study avoid inappropriate exclusions? • Unclear *The inclusion criteria are vague with regards to imaging or the standards/guidelines that were used.* RISK Could the selection of patients have introduced bias? • Unclear CONCERN Is there concern that the included patients do not match the review question? • Low Were the index test results interpreted without knowledge of the results of the reference standard? • Unclear *Blinding is not mentioned.* If a threshold was used, was it pre-specified? • Yes RISK Could the conduct or interpretation of the index test have introduced bias? • Unclear

Short Title	Title	Study Characteristics	Risk of Bias
		Downstream investigations and/or treatments • Recommendation of open thoracotomy or video-assisted thoracic surgery with systematic lymph node dissection to people whose endoscopic staging results did not show mediastinal masses **Protocol outcome measures** • Diagnostic accuracy *The diagnostic standard for a malignant result was the pathological confirmation of malignancy by any tissue sampling (EBUS-TBNA, EUS-FNA or surgical biopsy). The diagnostic standard for a benign result was the surgical confirmation of lesions showing no malignancy. The diagnostic accuracy, sensitivity and negative predictive value (NPV) for the detection of mediastinal metastasis (N2 or N3) were calculated using the standard definitions.* • Diagnostic sensitivity • Diagnostic negative predictive value • Safety: pneumothorax • Patient acceptability	Concerns regarding applicability • Low Is the reference standard likely to correctly classify the target condition? • Yes Were the reference standard results interpreted without knowledge of the results of the index test? • Unclear *Blinding is not mentioned* RISK Could the reference standard, its conduct, or its interpretation have introduced bias? • Low Was there an appropriate interval between index test(s) and reference standard? • Unclear *Timing is not mentioned* Did all patients receive a reference standard? • Yes Did patients receive the same reference standard? • Yes Were all patients included in the analysis?

Short Title	Title	Study Characteristics	Risk of Bias
			• Yes RISK Could the patient flow have introduced bias? • Low Overall quality • Moderate
Larsen 2005	Endoscopic ultrasound guided biopsy performed routinely in lung cancer staging spares futile thoracotomies: preliminary results from a randomised clinical trial	**Study type** • Randomised controlled trial **Study details** • Study location *Denmark* • Study setting *Gentofte University Hospital* • Study dates *November 2001 to February 2004* • Duration of follow-up *The median follow-up time from inclusion date was 1.3 years (range 0.2-2.4 years) in the routine EUS-FNA group and 1.4 years (range 0.2-2.4 years) in the group that had EUS-FNA only if CT showed invasion adjacent to the oesophagus* • Sources of funding *Not disclosed* • Lung cancer staging system used *American College of Chest Physicians. Lung cancer. Invasive staging: the guidelines. Chest 2003; 123: 167-175*	**Quality assessment (RCT)** Random sequence generation • Low risk of bias Allocation concealment • Low risk of bias Blinding of outcome assessment • Unclear risk of bias *Blinding of pathologists was not mentioned.* Blinding of participants and personnel • Unclear risk of bias *Not possible* Incomplete outcome data • Low risk of bias Selective reporting • Low risk of bias

Short Title	Title	Study Characteristics	Risk of Bias
		Inclusion criteria • Suspected or diagnosed lung cancer after CT/PET, bronchoscopy, TBNA/TTNA, lung function tests and general examination **Exclusion criteria** • <18 years of age • Not fit enough to undergo thoracotomy and lung resection • Pregnancy • Verified N2/3-, T4- or M1-disease or small-cell lung cancer **Sample characteristics** • Sample size *59 people* • Split between study groups *EUS-FNA for all = 28; EUS-FNA only if CT showed invasion adjacent to the oesophagus = 31* • Loss to follow-up *Three people in the EUS-FNA for all group did not undergo EUS-FNA because one became medically unfit, one person had had M1-disease (contra-lateral lung metastasis) verified before EUS-FNA was performed and one patient refused EUS-FNA on the day of examination.* • %female *EUS-FNA for all = 43% female, 57% male; EUS-FNA only if CT showed invasion adjacent to the oesophagus = 47% female, 53% male* • Mean age (SD) *EUS-FNA for all = 64 years (10); EUS-FNA only if CT showed invasion adjacent to the oesophagus = 65 years (10)*	Other sources of bias • Low risk of bias Overall risk of bias • Low **QUADAS 2** Was a random sample of patients enrolled? • Yes Was a case-control design avoided? • Yes Did the study avoid inappropriate exclusions? • Yes RISK Could the selection of patients have introduced bias? • Low CONCERN Is there concern that the included patients do not match the review question? • Low Were the index test results interpreted without knowledge of the results of the reference standard? • Unclear

Short Title	Title	Study Characteristics	Risk of Bias
		• Nodal staging on initial PET/CT scan *CT stage (I-V): EUS-FNA for all = IA: 9%; IB: 6%; IIB: 4%; IIIA: 19%; IIIB: 36%; IV: 26%; EUS-FNA only if CT showed invasion adjacent to the oesophagus = IA: 12%; IB: 4%; IIB: 6%; IIIA: 25%; IIIB: 35%; IV: 18%* **Interventions** • Mediastinoscopy + EUS-FNA for all • Mediastinoscopy + EUS-FNA only if CT showed invasion adjacent to the oesophagus **Downstream investigations and/or treatments** • Surgical resection or multimodal therapy *Provided mediastinal metastases were demonstrated by EUS-FNA, or if direct mediastinal organ invasion was demonstrated by EUS, in concordance with a CT suspicion, a malignant cytological diagnosis obtained by EUS-FNA was taken as final proof of malignancy in the mediastinum. The options for post-staging treatment of NSCLC, during the study period, were in general: 1) Surgical resection, provided no tumour-spread outside the lung was found; 2) Induction chemotherapy followed by resection in patients with ipsilateral mediastinal lymph node metastases (stage IIIA-N2); or 3) Chemo-/radiotherapy alone if contralateral mediastinal- or distant metastases were present (stage IIIB and IV).* **Protocol outcome measures** • Safety: other complications **Non-protocol outcome measures**	*Blinding of the pathologists was not mentioned.* If a threshold was used, was it pre-specified? • Yes RISK Could the conduct or interpretation of the index test have introduced bias? • Unclear Concerns regarding applicability • Low Is the reference standard likely to correctly classify the target condition? • Yes Were the reference standard results interpreted without knowledge of the results of the index test? • Unclear *Blinding of pathologists was not mentioned.* RISK Could the reference standard, its conduct, or its interpretation have introduced bias? • Low CONCERN Is there concern that the target condition as defined by the reference standard does not match the review question?

Short Title	Title	Study Characteristics	Risk of Bias
		• No. of avoidable thoracotomies *A thoracotomy was classified as futile/avoidable if: 1) An intended curative thoracotomy ended as an explorative thoracotomy without tumour resection; or 2) A resected patient died from lung cancer or had recurrent disease during follow up.* • Percentage (or number) of people who died during a specified follow-up period • Recurrence during a specified follow-up period	• Low Was there an appropriate interval between index test(s) and reference standard? • Unclear *Timing was not mentioned.* Did all patients receive a reference standard? • Yes Did patients receive the same reference standard? • Yes Were all patients included in the analysis? • Yes RISK Could the patient flow have introduced bias? • Low Overall quality • High
Navani 2015	Lung cancer diagnosis and staging with endobronchial ultrasound-guided transbronchial needle aspiration	**Study type** • Randomised controlled trial *They randomly assigned participants (1:1) to either conventional diagnosis and staging (CDS group) or EBUS-TBNA as an initial investigation after a staging CT scan followed by further diagnosis and staging techniques if needed (EBUS group). They used a telephone randomisation method with permuted computer-generated blocks of*	**Quality assessment (RCT)** Random sequence generation • Low risk of bias Allocation concealment • Low risk of bias

Short Title	Title	Study Characteristics	Risk of Bias
	compared with conventional approaches: an open-label, pragmatic, randomised controlled trial	*four. Randomisation was stratified according to the presence of mediastinal lymph nodes that measured 1 cm or more in the short axis and by recruiting centre. An investigator undertook the informed consent process, followed by the telephone randomisation process done by research assistants. The random allocation sequence was kept in the randomisation centre and concealed from participants and investigators until the interventions were assigned. Because of the nature of the intervention, masking of participants and consenting investigators was not possible. However, pathologists and radiologists were unaware that patients were enrolled into a clinical trial. Data were obtained on paper-based case forms and entered by an independent clerk onto a secured trial database on a dedicated trial computer.* **Study details** • Study location *UK* • Study setting *University College London Hospital, Whittington Hospital, North Middlesex University Hospital, Princess Alexandra Hospital, Barnet General Hospital, and Nottingham University Hospital* • Study dates *June 2008 to July 2011* • Duration of follow-up *Not stated. However, the survival curve has data collected for just over a 4-year duration. The final diagnosis of nodal staging was established in both groups by clinical follow-up of at least 1 year and pathological changes noted with EBUS-TBNA, conventional TBNA, EUS-FNA, mediastinoscopy, or dissection of mediastinal lymph nodes.* • Sources of funding *UK Medical Research Council*	Blinding of outcome assessment • Unclear risk of bias *Because of the nature of the intervention, masking of participants and consenting investigators was not possible. However, pathologists and radiologists were unaware that patients were enrolled into a clinical trial.* Blinding of participants and personnel • Unclear risk of bias *Because of the nature of the intervention, masking of participants and consenting investigators was not possible. However, pathologists and radiologists were unaware that patients were enrolled into a clinical trial.* Incomplete outcome data • Low risk of bias Selective reporting • Low risk of bias Other sources of bias • Low risk of bias Overall risk of bias • Low

Short Title	Title	Study Characteristics	Risk of Bias
		• Lung cancer staging system used *7th edition of the tumour, node, metastasis (TNM) staging system 2012* **Inclusion criteria** • Suspected stage I to IIIA lung cancer on CT neck, thorax and upper abdomen **Exclusion criteria** • <18 years of age • Not fit enough to undergo thoracotomy and lung resection • Significant concurrent malignant disease • Any condition that contraindicated the intervention or mediastinoscopy • Any medication that contraindicated the intervention or mediastinoscopy • Known extrathoracic malignant disease • Supraclavicular lymphadenopathy • Pleural effusion **Sample characteristics** • Sample size *132 people with suspected lung cancer* • Split between study groups *EBUS-TBNA / EUS-FNA = 66 people; CDS (Bronchoscopy / CT-guided biopsy) = 66 people* • Loss to follow-up	Directness • Directly applicable **QUADAS 2** Was a random sample of patients enrolled? • Yes Was a case-control design avoided? • Yes Did the study avoid inappropriate exclusions? • Yes RISK Could the selection of patients have introduced bias? • Low CONCERN Is there concern that the included patients do not match the review question? • Low Were the index test results interpreted without knowledge of the results of the reference standard? • Yes If a threshold was used, was it pre-specified? • Yes

Short Title	Title	Study Characteristics	Risk of Bias
		One patient (randomly assigned to CDS) declined all further investigations and withdrew consent before any investigations were done. • %female *EBUS-TBNA / EUS-FNA = 35% CDS (Bronchoscopy / CT-guided biopsy) = 30%* • Mean age (SD) *EBUS-TBNA / EUS-FNA = 71 years (IQR 62-78) CDS (Bronchoscopy / CT-guided biopsy) = 68 years (IQR 61-73)* • Smoking history *EBUS-TBNA / EUS-FNA = 28.1% CDS (Bronchoscopy / CT-guided biopsy) = 23.4%* • Nodal staging on initial PET/CT scan *EBUS-TBNA / EUS-FNA = N0: 32%; N1: 9%; N2: 51%; N3: 8%; CDS (Bronchoscopy / CT-guided biopsy) = N0: 30%; N1: 14%; N2: 50%; N3: 6%* **Interventions** • EBUS-TBNA as initial investigation. EUS-FNA if target node cannot be accessed by EBUS-TBNA *In the EBUS group, 64 (97%) of 66 underwent EBUS and two (3%) had EUS-FNA as an initial procedure. Five (8%) of 66 patients had a subsequent radiology-guided biopsy sample taken.* • Bronchoscopy or CT-guided biopsy (NHS conventional diagnosis and staging) *Participants allocated to conventional diagnosis and staging (CDS) underwent investigations as determined by the local multidisciplinary team. The investigators suggested an algorithm for CDS in the trial protocol based on the most recently available NICE guidance (2005) at*	RISK Could the conduct or interpretation of the index test have introduced bias? • Low Concerns regarding applicability • Low Is the reference standard likely to correctly classify the target condition? • Yes Were the reference standard results interpreted without knowledge of the results of the index test? • Unclear RISK Could the reference standard, its conduct, or its interpretation have introduced bias? • Low CONCERN Is there concern that the target condition as defined by the reference standard does not match the review question? • Low Was there an appropriate interval between index test(s) and reference standard? • Yes

Short Title	Title	Study Characteristics	Risk of Bias
		the time the trial started. The trial management group agreed that allowing the responsible multidisciplinary teams to determine the patients' investigations would provide the best comparator group. This allowed the control CDS group to emulate clinical practice, giving the trial strong external validity. In the CDS group, 44 (67%) of 66 patients initially underwent a bronchoscopy and 29 (44%) had a radiology-guided biopsy sample taken. 5 underwent conventional TBNA, 1 underwent a mediastinoscopy. 2 underwent a PET-CT scan. **Protocol outcome measures** • Diagnostic accuracy *Diagnostic accuracy percentages were included for the EBUS-TBNA/EUS-FNA arm but not for the conventional diagnosis and staging arm. Therefore, these numbers were excluded because our protocol's inclusion criteria are RCTs where the results of one arm are compared against the other.* • Safety: mortality • Safety: in-patient admission • Safety: pneumothorax • Safety: other complications • Timing: time to treatment decision *Time from first outpatient appointment with the respiratory specialist to treatment decision by the multidisciplinary team, after completion of the diagnosis and staging procedures.* • Timing: time to diagnosis and staging *Percentage of people who had diagnosis and staging completed by 14 days* • No. of investigations / person	Did all patients receive a reference standard? • Yes Did patients receive the same reference standard? • Yes Were all patients included in the analysis? • Yes RISK Could the patient flow have introduced bias? • Low Overall quality • High

Short Title	Title	Study Characteristics	Risk of Bias
		Non-protocol outcome measures • Proportion of people diagnosed and staged with one investigation • No. of avoidable thoracotomies *An avoidable thoracotomy was defined as an open and close procedure, unexpected mediastinal nodal metastases (pN2/pN3), pT4 or pM1a/b disease, resection of benign disease or disease recurrence, or death within 1 year of thoracotomy.* • Duration of survival (time) • Duration of survival (Hazard Ratio)	
Tournoy 2008	Endoscopic ultrasound reduces surgical mediastinal staging in lung cancer: a randomized trial. American Journal of Respiratory & Critical Care Medicine	**Study type** • Randomised controlled trial **Study details** • Study location *Belgium* • Study setting *Ghent University Hospital. EUS-FNA was performed in an outpatient setting* • Study dates *December 2005 to January 2007* • Duration of follow-up *Participants were followed up until discharge after the procedure (1 to 22 nights, with a median of 2 nights)* • Sources of funding *Not mentioned. The authors disclosed that they did not have a financial relationship with a commercial entity that had an interest in the study.* • Lung cancer staging system used	**Quality assessment (RCT)** Random sequence generation • Unclear risk of bias *Method not mentioned* Allocation concealment • Unclear risk of bias *Not mentioned* Blinding of outcome assessment • Unclear risk of bias *Not mentioned* Blinding of participants and personnel • Unclear risk of bias *Not possible* Incomplete outcome data

Short Title	Title	Study Characteristics	Risk of Bias
		Not stated. In the reference section, the following guidelines were referred to: Detterbeck FC, DeCamp MM Jr, Kohman LJ, Silvestri GA. Lung cancer: invasive staging: the guidelines. Chest 2003;123:167S–175S. Detterbeck FC, Jantz MA, Wallace MB, Vansteenkiste J, Silvestri GA; American College of Chest Physicians. Invasive mediastinal staging of lung cancer: ACCP evidence-based clinical practice guidelines, 2nd ed. Chest 2007;132:202S–220S. **Inclusion criteria** • Proven or suspected NSCLC • Suspected mediastinal lymph node invasion on CT/PET *Their guidelines for invasive mediastinal exploration were enlarged (>1-cm short axis) mediastinal lymph nodes and/or FDG uptake in the mediastinal lymph nodes, tumours abutting the mediastinum regardless of FDG uptake in the lymph nodes, and absence of FDG uptake in the primary tumour.* **Exclusion criteria** • Not fit enough to undergo thoracotomy and lung resection • Any condition that contraindicated the intervention or mediastinoscopy • Any medication that contraindicated the intervention or mediastinoscopy • Unresectable tumour • No distant metastasis • Former therapy for lung cancer • Concurrent other malignancy **Sample characteristics**	• Low risk of bias Selective reporting • Low risk of bias Other sources of bias • Low risk of bias Overall risk of bias • Moderate Directness • Directly applicable **QUADAS 2** Was a random sample of patients enrolled? • Unclear *Method not mentioned* Was a case-control design avoided? • Yes Did the study avoid inappropriate exclusions? • Yes RISK Could the selection of patients have introduced bias?

Short Title	Title	Study Characteristics	Risk of Bias
		• Sample size *40 people* • Split between study groups *EUS-FNA = 19; Straight to surgical staging = 21* • Loss to follow-up *None* • %female *EUS-FNA = 11% female, 89% male; Straight to surgical staging = 5% female, 95% male* • Mean age (SD) *EUS-FNA = 67 years (range 47-78); Straight to surgical staging = 61 years (range 42-74)* • Nodal staging on initial PET/CT scan *EUS-FNA = N2: 79%; N3: 21%; T1: 5%; T2: 84%; T3: 0%; T4: 11%; Straight to surgical staging = N2: 67%; N3: 33%; T1: 10%; T2: 76%; T3: 5%; T4: 10%* **Interventions** • Straight to surgical staging (mediastinoscopy) • Mediastinoscopy + EUS-FNA for all **Downstream investigations and/or treatments** • Surgical staging if required, then thoracotomy if required **Protocol outcome measures** • Diagnostic sensitivity • Diagnostic specificity • Diagnostic negative predictive value	• Low CONCERN Is there concern that the included patients do not match the review question? • Low Were the index test results interpreted without knowledge of the results of the reference standard? • Unclear If a threshold was used, was it pre-specified? • Yes RISK Could the conduct or interpretation of the index test have introduced bias? • Unclear Concerns regarding applicability • Low Is the reference standard likely to correctly classify the target condition? • Yes Were the reference standard results interpreted without knowledge of the results of the index test? • Unclear

Short Title	Title	Study Characteristics	Risk of Bias
		• Diagnostic positive predictive value • Safety: in-patient admission • Safety: other complications	RISK Could the reference standard, its conduct, or its interpretation have introduced bias? • Low CONCERN Is there concern that the target condition as defined by the reference standard does not match the review question? • Low Was there an appropriate interval between index test(s) and reference standard? • Unclear *Not mentioned* Did all patients receive a reference standard? • Yes Did patients receive the same reference standard? • Yes Were all patients included in the analysis? • Yes RISK Could the patient flow have introduced bias? • Low Overall quality • Moderate

Appendix F – GRADE tables

RQ 1.1: Mediastinoscopy + EUS-FNA vs mediastinoscopy + EUS-FNA only if CT shows invasion adjacent to the oesophagus: intervention evidence

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	EUS-FNA	EUS-FNA if CT shows invasion	Summary of results (95% CI)	
Safety: complications (RR >1 favours EUS-FNA if CT shows invasion adjacent to the oesophagus)									
1 (Larsen 2005)	RCT	Not serious	Not serious	N/A	Serious[1]	53	51	N/A[2]	Moderate
Safety: number of avoidable thoracotomies (RR >1 favours EUS-FNA if CT shows invasion adjacent to the oesophagus)									
1 (Larsen 2005)	RCT	Not serious	Not serious	N/A	Not serious	53	51	RR 0.37 (0.14, 0.96)	High
Recurrence or death during a median follow-up time of 1.3 years (range 0.2-2.4 years) for routine EUS-FNA and 1.4 years (range 0.2-2.4 years) for EUS-FNA if local invasion (RR >1 favours EUS-FNA if CT shows invasion adjacent to the oesophagus)									
1 (Larsen 2005)	RCT	Not serious	Not serious	N/A	Serious[1]	53	51	RR 0.48 (0.15, 1.50)	Moderate

1. Non-significant result
2. Not applicable - no events in either arm

RQ 1.1: EUS-FNA vs straight to surgical staging: intervention evidence

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	EUS-FNA	Straight to surgical staging	Summary of results (95% CI)	
Safety: in-patient admission for staging only, median number of nights									
1 (Tournoy 2008)	RCT	Not serious	Not serious	N/A	Serious[1]	19	21	EUS-FNA: median = 0 nights; straight to surgical staging: median = 2 nights (range: 1-22)[2]	Moderate
Safety: perforation / bleeding (RR >1 favours surgical staging)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	EUS-FNA	Straight to surgical staging	Summary of results (95% CI)	
1 (Tournoy 2008)	RCT	Not serious	Not serious	N/A	Very serious[1,3]	19	21	RR 0.37 (0.02, 8.50)	Low

1. Small number of participants. Downgraded once because the sample size is 26 to 40
2. These results are presented as they are because they are expressed as medians
3. Non-significant result

RQ 1.1: EUS-FNA vs straight to surgical staging: diagnostic accuracy evidence. Reference standards: For benign results, surgical confirmation. For malignant results, pathology

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Negative predictive value (95%CI)	Prevalence	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
EUS-FNA for all										
1 (Tournoy 2008)	RCT	19	93.0% (66.0, 99.0)	83.0% (35.0%, 99.0)	73.7%	Not serious	Not serious	N/A	Very serious[1]	Low
Straight to surgical staging										
1 (Tournoy 2008)	RCT	21	73.0% (39.0, 93.0)	73.0% (39.0, 93.0)	52.3%	Not serious	Not serious	N/A	Very serious[1]	Low

1. Very small number of participants. Downgraded twice because the sample size is below 25

RQ 1.1 and RQ 1.2: Bronchoscopy, EBUS-TBNA then EUS-FNA if necessary on mediastinal nodes inaccessible or difficult to access by EBUS-TBNA vs bronchoscopy, EUS-FNA then EBUS-TBNA if necessary on mediastinal nodes inaccessible or difficult to access by EUS-FNA: intervention evidence

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	EBUS-TBNA then EUS-FNA	EUS-FNA then EBUS-TBNA	Summary of results (95% CI)	
Safety: pneumothorax (RR >1 favours EUS-FNA then EBUS-TBNA)									
1 (Kang 2014)	RCT	Serious[1]	Serious[2]	N/A	Serious[3]	80	80	RR 0.33 (0.01, 8.20)	Very low
Patient satisfaction: overall tolerance at 3-5 days after the interventions. Visual analogue scale from 1-10 (values >0 EUS-FNA then EBUS-TBNA)									
1 (Kang 2014)	RCT	Serious[1]	Serious[2]	N/A	Serious[3]	80	80	MD -0.54 (-1.28, 0.20)	Very low

1. Vague inclusion criteria
2. Both arms of the trial involve giving patients 3 endoscopic interventions. Therefore, this is indirect evidence because in the UK, healthcare professionals aim to use fewer endoscopic interventions
3. Non-significant result

RQ 1.1 and RQ 1.2: Bronchoscopy, EBUS-TBNA then EUS-FNA if necessary vs bronchoscopy, EUS-FNA then EBUS-TBNA if necessary: diagnostic accuracy evidence. Reference standards: For benign results, surgical confirmation. For malignant results, pathology

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Negative predictive value (95%CI)	Prevalence	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
Bronchoscopy, EBUS-TBNA, then EUS-FNA arm										
1 (Kang 2014)	RCT	74	85.3% (68.3, 93.0)	88.0% (75.1, 94.7)	45.9%	Serious[1]	Serious[2]	N/A	Not serious	Low
Bronchoscopy, EUS-FNA, then EBUS-TBNA arm										
1 (Kang 2014)	RCT	74	90.4% (71.8, 97.2)	95.2% (84.8, 98.6)	33.8%	Serious[1]	Serious[2]	N/A	Not serious	Low
Bronchoscopy, EBUS-TBNA, then EUS-FNA arm: EBUS-TBNA only										

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Negative predictive value (95%CI)	Prevalence	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
1 (Kang 2014)	RCT	74	81.4% (65.2, 91.1)	86.2% (73.1, 93.4)	45.9%	Serious[1]	Serious[2]	N/A	Not serious	Low
Bronchoscopy, EUS-FNA, then EBUS-TBNA arm: EUS-FNA only										
1 (Kang 2014)	RCT	74	59.6% (40.3, 76.4)	82.5% (70.8, 90.2)	33.8%	Serious[1]	Serious[2]	N/A	Not serious	Low

1. Vague inclusion criteria
2. Both arms of the trial involve giving patients 3 endoscopic interventions. Therefore, this is indirect evidence because in the UK, healthcare professionals aim to use fewer endoscopic interventions

RQ 1.1 and RQ 1.2: EBUS-TBNA (or EUS-FNA) vs conventional (bronchoscopy or CT-guided biopsy etc): intervention evidence

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	EBUS-TBNA (or EUS-FNA)	Conventional	Summary of results (95% CI)	
Safety: pneumothorax (RR >1 favours conventional (bronchoscopy or CT-guided biopsy etc))									
1 (Navani 2015)	RCT	Not serious	Not serious	N/A	Serious[1]	66	66	RR 1.00 (0.06, 15.65)	Moderate
Safety: in-patient admissions (RR >1 favours conventional (bronchoscopy or CT-guided biopsy etc))									
1 (Navani 2015)	RCT	Not serious	Not serious	N/A	Serious[1]	66	66	RR 0.33 (0.01, 8.04)	Moderate
Timing: time to treatment decision									
1 (Navani 2015)	RCT	Not serious	Not serious	N/A	Not serious	66	66	EBUS-TBNA/EUS-FNA: median = 14 days (14-15); bronchoscopy = 29 days (23-35) [2]	High
Timing: number of people who had diagnosis and staging completed by 14 days (RR >1 favours EBUS-TBNA (or EUS-FNA))									
1 (Navani 2015)	RCT	Not serious	Not serious	N/A	Not serious	66	66	RR 4.38 (2.20, 8.71)	High
Number of investigations per person (values >0 favour conventional (bronchoscopy or CT-guided biopsy etc))									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	EBUS-TBNA (or EUS-FNA)	Conventional	Summary of results (95% CI)	
1 (Navani 2015)	RCT	Not serious	Not serious	N/A	Not serious	66	66	MD -0.69 (-0.95, -0.43)	High
Number of people diagnosed and staged with one investigation (RR >1 favours EBUS-TBNA (or EUS-FNA))									
1 (Navani 2015)	RCT	Not serious	Not serious	N/A	Not serious	66	66	RR 3.75 (1.86, 7.56)	High
Number of avoidable thoracotomies at 1 year (RR >1 favours EBUS-TBNA (or EUS-FNA))									
1 (Navani 2015)	RCT	Not serious	Not serious	N/A	Serious[1]	66	66	RR 2.60 (0.98, 6.88)	Moderate
Duration of survival: median number of days									
1 (Navani 2015)	RCT	Not serious	Not serious	N/A	Serious[1]	66	66	EBUS-TBNA/EUS-FNA: median = 503 days (312-715); bronchoscopy = 312 days (231-488) [2]	Moderate
Duration of survival: hazard ratio (HR >1 favours conventional (bronchoscopy / CT guided biopsy etc))									
1 (Navani 2015)	RCT	Not serious	Not serious	N/A	Not serious	66	66	HR 0.60 (0.37, 0.98)	High

1. Non-significant result
2. These results are presented as they are because they are expressed as medians

RQ 1.1 and RQ 1.2: EBUS-TBNA (or EUS-FNA) vs conventional (bronchoscopy or CT-guided biopsy etc): diagnostic accuracy evidence. Reference standards: For benign results, surgical confirmation. For malignant results, pathology

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Negative predictive value (95%CI)	Prevalence	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
EBUS-TBNA. If node cannot be accessed, then EUS-FNA										
1 (Navani 2015)	RCT	66	92.0% (78.0, 98.0)	90.0% (72.0, 97.0)	75.8%	Not serous	Not serious	N/A	Not serious	High

RQ 1.2: EUS-FNA followed by EBUS-TBNA vs straight to surgical staging: intervention evidence

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	EUS-FNA followed by EBUS-TBNA	Straight to surgical staging	Summary of results (95% CI)	
Safety: pneumothorax (RR >1 favours surgical staging)									
1 (Annema 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	123	118	RR 0.96 (0.06, 15.16)	Low
Safety: total number of complications (RR >1 favours surgical staging)									
1 (Annema 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	123	118	RR 0.82 (0.28, 2.38)	Low
Quality of life change at 6 months from randomisation, EQ-5D (values >0 favour EUS-FNA + EBUS-TBNA)									
1 (Annema 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	123	118	MD 0.01 (-0.07, 0.09)	Low
Number of avoidable thoracotomies (RR >1 favours surgical staging)									
1 (Annema 2010)	RCT	Serious[1]	Not serious	N/A	Not serious	123	118	RR 0.41 (0.20, 0.86)	Moderate
Number of people who died between staging and 6 months later (RR >1 favours surgical staging)									
1 (Annema 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	123	118	RR 0.78 (0.34, 1.83)	Low

1. Details of randomisation not given
2. Non-significant result

RQ 1.2: EUS-FNA followed by EBUS-TBNA vs straight to surgical staging: diagnostic accuracy evidence. Reference standards: For benign results, surgical confirmation. For malignant results, pathology

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Negative predictive value (95%CI)	Prevalence	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
EUS-FNA followed by EBUS-TBNA										

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Negative predictive value (95%CI)	Prevalence	Risk of bias	Indirectness	Inconsistency	Imprecision	Quality
1 (Annema 2010)	RCT	123	93.3% (84.2, 97.3)	92.7% (83.0, 97.1)	53.7%	Serious[1]	Not serious	N/A	Not serious	Moderate
Straight to surgical staging (mediastinoscopy)										
1 (Annema 2010)	RCT	117	78.3% (65.3, 87.4)	85.3% (75.6, 91.5%)	44.1%	Serious[1]	Not serious	N/A	Not serious	Moderate

1. Details of randomisation not given

1 # Appendix G – Excluded Studies

2 ## Excluded clinical studies

3

Short title	Title	Reason for exclusion
Adams (2009)	Test performance of endobronchial ultrasound and transbronchial needle aspiration biopsy for mediastinal staging in patients with lung cancer: systematic review and meta-analysis	Systematic review of non-randomised controlled trials
Akulian (2014)	Molecular profiling of adenocarcinoma of the lung	Review article but not a systematic review
Almeida (2012)	Bronchoscopy and endobronchial ultrasound for diagnosis and staging of lung cancer	Review article but not a systematic review
Anantham (2010)	Endobronchial ultrasound-guided transbronchial needle aspiration in the diagnosis and staging of lung cancer	Review article but not a systematic review
Boonsarngsuk (2015)	Comparison of diagnostic performances among bronchoscopic sampling techniques in the diagnosis of peripheral pulmonary lesions	Non-randomised study
Casal (2012)	Randomized clinical trial of endobronchial ultrasound needle biopsy with and without aspiration	No relevant outcomes. The randomisation is not between two different arms of a trial. Lung cancer is mentioned as a coincidence, it is not the main focus
Chao 2009	Endobronchial ultrasonography-guided transbronchial needle aspiration increases the diagnostic yield of peripheral pulmonary lesions: a randomized trial	This study is on radial EBUS, which is not in the protocol
Dango (2010)	Endobronchial ultrasound-guided transbronchial needle aspiration and its role in non-small cell lung cancer: Diagnostic impact and limitations	Review article but not a systematic review
Darwiche (2013)	Assessment of SHOX2 methylation in EBUS-TBNA specimen improves accuracy in lung cancer staging	Non-randomised study
Ernst (2008)	Diagnosis of mediastinal adenopathy-real-time endobronchial ultrasound guided needle aspiration versus mediastinoscopy	Non-randomised study
Fritscher-Ravens (2003)	Mediastinal lymph node involvement in potentially resectable lung cancer: comparison of CT, positron emission tomography, and endoscopic ultrasonography with and without fine-needle aspiration	Non-randomised study

Short title	Title	Reason for exclusion
Fritscher-Ravens (2003)	Endoscopic ultrasound evaluation in the diagnosis and staging of lung cancer	Review article but not a systematic review
Godbout (2016)	Evaluation of pulmonary nodules using the spyglass direct visualization system combined with radial endobronchial ultrasound: A clinical feasibility study	Non-randomised study
Gompelmann (2014)	Role of endobronchial and endoscopic ultrasound in pulmonary medicine	Review article but not a systematic review
Govert (1999)	A prospective comparison of fiberoptic transbronchial needle aspiration and bronchial biopsy for bronchoscopically visible lung carcinoma	Non-randomised study
Grah (2011)	Comparison of 21 gauge and 22-gauge aspiration needle during endobronchial ultrasound-guided transbronchial needle aspiration: a randomised trial	Conference abstract
Gu (2009)	Endobronchial ultrasound-guided transbronchial needle aspiration for staging of lung cancer: a systematic review and meta-analysis	Systematic review of non-randomised controlled trials
Hassan (2010)	Comparative study of efficacy of brush cytology and transthoracic fine needle aspiration cytology in the diagnosis of bronchogenic carcinoma	Non-randomised study
Herth (2004)	Conventional vs Endobronchial Ultrasound-Guided Transbronchial Needle Aspiration: A Randomized Trial	No relevant outcomes. The outcome of interest is diagnostic yield. Diagnostic yield is the likelihood that a test or procedure will provide the information needed to establish a diagnosis. It is not a measurement of diagnostic accuracy.
Herth (2005)	Transbronchial versus transesophageal ultrasound-guided aspiration of enlarged mediastinal lymph nodes	Non-randomised study
Hwangbo (2010)	Transbronchial and transesophageal fine-needle aspiration using an ultrasound bronchoscope in mediastinal staging of potentially operable lung cancer	Non-randomised study
Jiang (2014)	TBNA with and without EBUS: A comparative efficacy study for the diagnosis and staging of lung cancer	Non-randomised study
Kramer (2003)	Current Concepts in the Mediastinal Lymph Node Staging of Nonsmall Cell Lung Cancer	Systematic review of non-randomised controlled trials
Lardinois (2011)	Pre- and intra-operative mediastinal staging in non-small-cell lung cancer	Review article but not a systematic review
Micames (2007)	Endoscopic ultrasound-guided fine-needle aspiration for non-small cell lung cancer staging: A systematic review and metaanalysis	Systematic review of non-randomised controlled trials

Short title	Title	Reason for exclusion
Mullan (2004)	CT-guided fine-needle aspiration of lung nodules: effect on outcome of using coaxial technique and immediate cytological evaluation	Non-randomised study
Oezkan (2017)	Feasibility study of using 19G needle for EBUS-TBNA: a prospective-randomized comparison of 19G and 22G EBUS-needles	Conference abstract
Ost (2016)	Diagnostic Yield and Complications of Bronchoscopy for Peripheral Lung Lesions. Results of the AQuIRE Registry	Non-randomised study
Paone (2005)	Endobronchial ultrasound-driven biopsy in the diagnosis of peripheral lung lesions	Study is on EBUS-TBB, not EBUS-TBNA
Puri (2009)	Randomized controlled trial of endoscopic ultrasound-guided fine-needle sampling with or without suction for better cytological diagnosis	No relevant outcomes. Lung cancer is mentioned as a coincidence, it is not the main focus
Roth 2011	A randomised trial of endobronchial ultrasound guided sampling in peripheral lung lesions	This study is on radial EBUS, not EBUS-TBNA
Saji (2011)	Comparison of 21-gauge and 22-gauge Needles for Endobronchial Ultrasound-Guided Transbronchial Needle Aspiration of Mediastinal and Hilar Lymph Nodes	Non-randomised study
Schreiber (2003)	Performance characteristics of different modalities for diagnosis of suspected lung cancer: Summary of published evidence	Systematic review of non-randomised controlled trials
Soja (2010)	Usefulness of transbronchial needle aspiration for initial lung cancer staging	Non-randomised study
Szlubowski (2012)	A comparison of the combined ultrasound of the mediastinum by use of a single ultrasound bronchoscope versus ultrasound bronchoscope plus ultrasound gastroscope in lung cancer staging: a prospective trial	Non-randomised study
Trisolini 2015	Randomized Trial of Endobronchial Ultrasound-Guided Transbronchial Needle Aspiration With and Without Rapid On-site Evaluation for Lung Cancer Genotyping	The comparison of EBUS-TBNA vs EBUS-TBNA with Rapid On-Site Evaluation (ROSE) is not in the protocol
Wagner (1989)	Transbronchial fine-needle aspiration. Reliability and limitations	Non-randomised study
Xi (2017)	Distant metastasis and survival outcomes after computed tomography-guided needle biopsy in resected stage I-III non-small cell lung cancer	Non-randomised study
Yarmus (2011)	A randomized prospective trial of the utility of rapid on-site evaluation of transbronchial needle aspirate specimens	Study on bronchoscopy

Short title	Title	Reason for exclusion
Yarmus (2015)	A randomized controlled trial evaluating airway inspection effectiveness during endobronchial ultrasound bronchoscopy	No relevant outcomes
Yasuda (2009)	Mediastinal lymph node staging in potentially resectable non-small cell lung cancer: a prospective comparison of CT and EUS/EUS-FNA	Non-randomised study
Zhang (2013)	Combined endobronchial and endoscopic ultrasound-guided fine needle aspiration for mediastinal lymph node staging of lung cancer: a meta-analysis	Systematic review of non-randomised controlled trials. There was one RCT included, which we are already including.

4 .

5 **Excluded economic studies**

Paper	Primary reason for exclusion
Bongers, M.L., Coupé, V.M., De Ruysscher, D., Oberije, C., Lambin, P. and Uyl-de Groot, C.A., 2015. Individualized Positron Emission Tomography–Based Isotoxic Accelerated Radiation Therapy Is Cost-Effective Compared With Conventional Radiation Therapy: A Model-Based Evaluation. *International Journal of Radiation Oncology* Biology* Physics*, *91*(4), pp.857-865.	Not conducted in a health care system similar to the UK.
Czarnecka-Kujawa, K., Rochau, U., Siebert, U., Atenafu, E., Darling, G., Waddell, T.K., Pierre, A., De Perrot, M., Cypel, M., Keshavjee, S. and Yasufuku, K., 2017. Cost-effectiveness of mediastinal lymph node staging in non–small cell lung cancer. *The Journal of thoracic and cardiovascular surgery*, *153*(6), pp.1567-1578.	Not conducted in a health care system similar to the UK.
Deppen, S.A., Davis, W.T., Green, E.A., Rickman, O., Aldrich, M.C., Fletcher, S., Putnam Jr, J.B. and Grogan, E.L., 2014. Cost-effectiveness of initial diagnostic strategies for pulmonary nodules presenting to thoracic surgeons. *The Annals of thoracic surgery*, *98*(4), pp.1214-1222.	Not conducted in a health care system similar to the UK.
Dietlein, M., Weber, K., Gandjour, A., Moka, D., Theissen, P., Lauterbach, K.W. and Schicha, H., 2000. Cost-effectiveness of FDG-PET for the management of potentially operable non-small cell lung cancer: priority for a PET-based strategy after nodal-negative CT results. *European journal of nuclear medicine*, *27*(11), pp.1598-1609.	Not conducted in a health care system similar to the UK.
Dietlein, M., Weber, K., Gandjour, A., Moka, D., Theissen, P., Lauterbach, K.W. and Schicha, H., 2000. Cost-effectiveness of FDG-PET for the management of solitary pulmonary nodules: a decision analysis based on cost reimbursement in Germany. *European journal of nuclear medicine*, *27*(10), pp.1441-1456.	Not conducted in a health care system similar to the UK.
Esnaola, N.F., Lazarides, S.N., Mentzer, S.J. and Kuntz, K.M., 2002. Outcomes and Cost-Effectiveness of Alternative Staging Strategies for Non–Small-Cell Lung Cancer. *Journal of clinical oncology*, *20*(1), pp.263-273.	Not conducted in a health care system similar to the UK.
Han, Y., Xiao, H., Zhou, Z., Yuan, M., Zeng, Y., Wu, H. and Fang, Y., 2015. Cost-effectiveness analysis of strategies introducing integrated 18F-FDG PET/CT into the mediastinal lymph node staging of non-small-cell lung cancer. *Nuclear medicine communications*, *36*(3), pp.234-241.	Not conducted in a health care system similar to the UK.
Hayashi, K., Abe, K., Yano, F., Watanabe, S., Iwasaki, Y. and Kosuda, S., 2005. Should mediastinoscopy actually be incorporated into the FDG PET strategy for patients with non-small cell lung carcinoma?. *Annals of nuclear medicine*, *19*(5), pp.393-398.	Not conducted in a health care system similar to the UK.

Paper	Primary reason for exclusion
Lejeune, C., Al Zahouri, K., Woronoff-Lemsi, M.C., Arveux, P., Bernard, A., Binquet, C. and Guillemin, F., 2005. Use of a decision analysis model to assess the medicoeconomic implications of FDG PET imaging in diagnosing a solitary pulmonary nodule. *The European Journal of Health Economics*, *6*(3), pp.203-214.	Not conducted in a health care system similar to the UK.
León, N.G., Escalona, S., Bandrés, B., Belda, C., Callejo, D. and Blasco, J.A., 2014. 18f-fluorodeoxyglucose positron emission tomography/computed tomography accuracy in the staging of non-small cell lung cancer: Review and cost-effectiveness. *Radiology research and practice*, *2014*.	Not conducted in a health care system similar to the UK.
Meyers, B.F., Haddad, F., Siegel, B.A., Zoole, J.B., Battafarano, R.J., Veeramachaneni, N., Cooper, J.D. and Patterson, G.A., 2006. Cost-effectiveness of routine mediastinoscopy in computed tomography–and positron emission tomography–screened patients with stage I lung cancer. *The Journal of thoracic and cardiovascular surgery*, *131*(4), pp.822-829.	Not conducted in a health care system similar to the UK.
Navani, N. and Janes, S.M., 2013. Endobronchial Ultrasound–guided Transbronchial Needle Aspiration for Lymphoma: The Final Frontier.	Not conducted in a health care system similar to the UK.
Navani, N., Nankivell, M., Woolhouse, I., Harrison, R.N., Munavvar, M., Oltmanns, U., Falzon, M., Kocjan, G., Rintoul, R.C. and Janes, S.M., 2011. Endobronchial ultrasound-guided transbronchial needle aspiration for the diagnosis of intrathoracic lymphadenopathy in patients with extrathoracic malignancy: a multicenter study. *Journal of Thoracic Oncology*, *6*(9), pp.1505-1509.	Not conducted in a health care system similar to the UK.
Navani, N., Lawrence, D.R., Kolvekar, S., Hayward, M., McAsey, D., Kocjan, G., Falzon, M., Capitanio, A., Shaw, P., Morris, S. and Omar, R.Z., 2012. Endobronchial ultrasound–guided transbronchial needle aspiration prevents mediastinoscopies in the diagnosis of isolated mediastinal lymphadenopathy: a prospective trial. *American journal of respiratory and critical care medicine*, *186*(3), pp.255-260.	Not conducted in a health care system similar to the UK.
Rintoul, R.C., Glover, M.J., Jackson, C., Hughes, V., Tournoy, K.G., Dooms, C., Annema, J.T. and Sharples, L.D., 2014. Cost effectiveness of endosonography versus surgical staging in potentially resectable lung cancer: a health economics analysis of the ASTER trial from a European perspective. *Thorax*, *69*(7), pp.679-681.	Not conducted in a health care system similar to the UK.
Sari, A.A., Ravaghi, H., Mobinizadeh, M. and Sarvari, S., 2013. The cost-utility analysis of PET-scan in diagnosis and treatment of non-small cell lung carcinoma in Iran. *Iranian Journal of Radiology*, *10*(2), p.61.	Not conducted in a health care system similar to the UK.
Schreyögg, J., Weller, J., Stargardt, T., Herrmann, K., Bluemel, C., Dechow, T., Glatting, G., Krause, B.J., Mottaghy, F., Reske, S.N. and Buck, A.K., 2010. Cost-effectiveness of hybrid PET/CT for staging of non-small cell lung cancer. *J Nucl Med*, *51*(11), pp.1668-75.	Not conducted in a health care system similar to the UK.
Scott, W.J., Shepherd, J. and Gambhir, S.S., 1998. Cost-effectiveness of FDG-PET for staging non–small cell lung cancer: a decision analysis. *The Annals of thoracic surgery*, *66*(6), pp.1876-1884.	Not conducted in a health care system similar to the UK.
Søgaard, R., Fischer, B.M.B., Mortensen, J., Rasmussen, T.R. and Lassen, U., 2013. The Optimality of Different Strategies for Supplemental Staging of Non–Small-Cell Lung Cancer: A Health Economic Decision Analysis. *Value in health*, *16*(1), pp.57-65.	Not conducted in a health care system similar to the UK.
van Loon, J., Grutters, J.P., Wanders, R., Boersma, L., Dingemans, A.M.C., Bootsma, G., Geraedts, W., Pitz, C., Simons, J., Brans, B. and Snoep, G.,	Not conducted in a health care system similar to the UK.

Paper	Primary reason for exclusion
2010. 18FDG-PET-CT in the follow-up of non-small cell lung cancer patients after radical radiotherapy with or without chemotherapy: an economic evaluation. *European Journal of Cancer*, *46*(1), pp.110-119.	
Wang, Y.T. and Huang, G., 2012. Is FDG PET/CT cost-effective for pre-operation staging of potentially operative non-small cell lung cancer?–from Chinese healthcare system perspective. *European journal of radiology*, *81*(8), pp.e903-e909.	Not conducted in a health care system similar to the UK.

Appendix H – References

Clinical Studies - Included

Annema J T, van Meerbeeck , J P, Rintoul R C, Dooms C, Deschepper E, Dekkers O M, De Leyn , P , Braun J, Carroll N R, Praet M, de Ryck , F , Vansteenkiste J, Vermassen F, Versteegh M I, Veselic M, Nicholson A G, Rabe K F, and Tournoy K G (2010) Mediastinoscopy vs endosonography for mediastinal nodal staging of lung cancer: a randomized trial. JAMA 304(20), 2245-52

Kang H J, Hwangbo B, Lee G K, Nam B H, Lee H S, Kim M S, Lee J M, Zo J I, Lee H S, and Han J Y (2014) EBUS-centred versus EUS-centred mediastinal staging in lung cancer: a randomised controlled trial. Thorax 69(3), 261-8

Larsen S S, Vilmann P, Krasnik M, Dirksen A, Clementsen P, Maltbaek N, Lassen U, Skov B G, and Jacobsen G K (2005) Endoscopic ultrasound guided biopsy performed routinely in lung cancer staging spares futile thoracotomies: preliminary results from a randomised clinical trial. Lung Cancer 49(3), 377-85

Navani N, Nankivell M, Lawrence D R, Lock S, Makker H, Baldwin D R, Stephens R J, Parmar M K, Spiro S G, Morris S, Janes S M, and Lung Boost trial investigators (2015) Lung cancer diagnosis and staging with endobronchial ultrasound-guided transbronchial needle aspiration compared with conventional approaches: an open-label, pragmatic, randomised controlled trial. The Lancet Respiratory Medicine 3(4), 282-9

Sharples L D, Jackson C, Wheaton E, Griffith G, Annema J T, Dooms C, Tournoy K G, Deschepper E, Hughes V, Magee L, Buxton M, Rintoul R C (2012) Clinical effectiveness and cost-effectiveness of endobronchial and endoscopic ultrasound relative to surgical staging in potentially resectable lung cancer: results from the ASTER randomised controlled trial. Health Technology Assessment 16(18), 1-100

Tournoy K G, De Ryck , F , Vanwalleghem L R, Vermassen F, Praet M, Aerts J G, Van Maele , G , van Meerbeeck , and J P (2008) Endoscopic ultrasound reduces surgical mediastinal staging in lung cancer: a randomized trial. American Journal of Respiratory & Critical Care Medicine 177(5), 531-5

Clinical studies – Excluded

Adams K, Shah P L, Edmonds L, and Lim E (2009) Test performance of endobronchial ultrasound and transbronchial needle aspiration biopsy for mediastinal staging in patients with lung cancer: systematic review and meta-analysis. Thorax 64(9), 757-62

Akulian J A, Belanger A R, and Yarmus L (2014) Molecular profiling of adenocarcinoma of the lung. Minerva Pneumologica 53(1), 1-8

Almeida F A (2012) Bronchoscopy and endobronchial ultrasound for diagnosis and staging of lung cancer. Cleveland Clinic Journal of Medicine 79 Electronic Suppl 1, eS11-6

Anantham D, and Koh M S (2010) Endobronchial ultrasound-guided transbronchial needle aspiration in the diagnosis and staging of lung cancer. Thoracic Cancer 1(1), 9-16

Boonsarngsuk V, Kanoksil W, and Laungdamerongchai S (2015) Comparison of diagnostic performances among bronchoscopic sampling techniques in the diagnosis of peripheral pulmonary lesions. Journal of Thoracic Disease 7(4), 697-703

Casal R F, Staerkel G A, Ost D, Almeida F A, Uzbeck M H, Eapen G A, Jimenez C A, Nogueras-Gonzalez G M, Sarkiss M, and Morice R C (2012) Randomized clinical trial of endobronchial ultrasound needle biopsy with and without aspiration. Chest 142(3), 568-573

Chao T Y, Chien M T, Lie C H, Chung Y H, Wang J L, and Lin M C (2009) Endobronchial ultrasonography-guided transbronchial needle aspiration increases the diagnostic yield of peripheral pulmonary lesions: a randomized trial. Chest 136(1), 229-236

Dango S, Guenter J, and Passlick B (2010) Endobronchial ultrasound-guided transbronchial needle aspiration and its role in non-small cell lung cancer: Diagnostic impact and limitations. Thoracic Cancer 1(2), 70-76

Darwiche K, Zarogoulidis P, Baehner K, Welter S, Tetzner R, Wohlschlaeger J, Theegarten D, Nakajima T, and Freitag L (2013) Assessment of SHOX2 methylation in EBUS-TBNA specimen improves accuracy in lung cancer staging. Annals of Oncology 24(11), 2866-70

Ernst A, Anantham D, Eberhardt R, Krasnik M, and Herth F J (2008) Diagnosis of mediastinal adenopathy-real-time endobronchial ultrasound guided needle aspiration versus mediastinoscopy. Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer 3(6), 577-82

Fritscher-Ravens A (2003) Endoscopic ultrasound evaluation in the diagnosis and staging of lung cancer. Lung Cancer 41(3), 259-67

Fritscher-Ravens A, Bohuslavizki K H, Brandt L, Bobrowski C, Lund C, Knofel W T, and Pforte A (2003) Mediastinal lymph node involvement in potentially resectable lung cancer: comparison of CT, positron emission tomography, and endoscopic ultrasonography with and without fine-needle aspiration. Chest 123(2), 442-51

Godbout K, Martel S, Simon M, Lampron N, and Delage A (2016) Evaluation of pulmonary nodules using the spyglass direct visualization system combined with radial endobronchial ultrasound: A clinical feasibility study. Open Respiratory Medicine Journal 10, 79-85

Gompelmann D, and Herth F J (2014) Role of endobronchial and endoscopic ultrasound in pulmonary medicine. Respiration 87(1), 3-8

Govert J A, Dodd L G, Kussin P S, and Samuelson W M (1999) A prospective comparison of fiberoptic transbronchial needle aspiration and bronchial biopsy for bronchoscopically visible lung carcinoma. Cancer 87(3), 129-34

Grah C, Griff S, Mairinger T, Bollmann L, Temme H, Seiberth R, and Vogelbusch M (2011) Comparison of 21 gauge and 22-gauge aspiration needle during endobronchial ultrasound-guided transbronchial needle aspiration: a randomised trial. Journal of thoracic oncology. 6(6 suppl. 2), S1148

Gu P, Zhao Y Z, Jiang L Y, Zhang W, Xin Y, and Han B H (2009) Endobronchial ultrasound-guided transbronchial needle aspiration for staging of lung cancer: a systematic review and meta-analysis. European Journal of Cancer 45(8), 1389-96

Hassan M Q, Ahamad M S, Ahmed S, and Chowdhury M A (2010) Comparative study of efficacy of brush cytology and transthoracic fine needle aspiration cytology in the diagnosis of bronchogenic carcinoma. Bangladesh Medical Research Council Bulletin 36(1), 37-9

Herth F, Becker H D, and Ernst A (2004) Conventional vs Endobronchial Ultrasound-Guided Transbronchial Needle Aspiration: A Randomized Trial. Chest 125(1), 322-325

Herth F J. F, Lunn W, Eberhardt R, Becker H D, and Ernst A (2005) Transbronchial versus transesophageal ultrasound-guided aspiration of enlarged mediastinal lymph nodes. American Journal of Respiratory and Critical Care Medicine 171(10), 1164-1167

Hwangbo B, Lee G K, Lee H S, Lim K Y, Lee S H, Kim H Y, Lee H S, Kim M S, Lee J M, Nam B H, and Zo J I (2010) Transbronchial and transesophageal fine-needle aspiration using an ultrasound bronchoscope in mediastinal staging of potentially operable lung cancer. Chest 138(4), 795-802

Jiang J, Browning R, Lechtzin N, Huang J, Terry P, and Wang K P (2014) TBNA with and without EBUS: A comparative efficacy study for the diagnosis and staging of lung cancer. Journal of Thoracic Disease 6(5), 416-420

Kramer H, and Groen H J. M (2003) Current Concepts in the Mediastinal Lymph Node Staging of Nonsmall Cell Lung Cancer. Annals of Surgery 238(2), 180-188

Lardinois D (2011) Pre- and intra-operative mediastinal staging in non-small-cell lung cancer. Swiss Medical Weekly 141(MARCH),

Micames C G, McCrory D C, Pavey D A, Jowell P S, and Gress F G (2007) Endoscopic ultrasound-guided fine-needle aspiration for non-small cell lung cancer staging: A systematic review and metaanalysis. Chest 131(2), 539-48

Mullan C P, Kelly B E, Ellis P K, Hughes S, Anderson N, and McCluggage W G (2004) CT-guided fine-needle aspiration of lung nodules: effect on outcome of using coaxial technique and immediate cytological evaluation. Ulster Medical Journal 73(1), 32-6

Oezkan F, Wolters C, Franzen D, Hager T, Eisenmann S, Freitag L, He K, Weinreich G, and Darwiche K (2017) Feasibility study of using 19G needle for EBUS-TBNA: a prospective-randomized comparison of 19G and 22G EBUS-needles. American journal of respiratory and critical care medicine. Conference: american thoracic society international conference, and ATS 2017. United states 195(no pagination),

Ost D E, Ernst A, Lei X, Kovitz K L, Benzaquen S, Diaz-Mendoza J, Greenhill S, Toth J, Feller-Kopman D, Puchalski J, Baram D, Karunakara R, Jimenez C A, Filner J J, Morice R C, Eapen G A, Michaud G C, Estrada Y Martin R. M, Rafeq S, Grosu H B, Ray C, Gilbert C R, Yarmus L B, Simoff M, and Registry A QuIRE Bronchoscopy (2016) Diagnostic Yield and Complications of Bronchoscopy for Peripheral Lung Lesions. Results of the AQuIRE Registry. American Journal of Respiratory & Critical Care Medicine 193(1), 68-77

Paone G, Nicastri E, Lucantoni G, Della Iacono, R , Battistoni P, D'Angeli A L, and Galluccio G (2005) Endobronchial ultrasound-driven biopsy in the diagnosis of peripheral lung lesions. Chest 128(5), 3551-3557

Puri R, Vilmann P, Saftoiu A, Skov B G, Linnemann D, Hassan H, Garcia E S. G, and Gorunescu F (2009) Randomized controlled trial of endoscopic ultrasound-guided fine-needle sampling with or without suction for better cytological diagnosis. Scandinavian Journal of Gastroenterology 44(4), 499-504

Roth K, Eagan T M, Andreassen A H, Leh F, and Hardie J A (2011) A randomised trial of endobronchial ultrasound guided sampling in peripheral lung lesions. Lung Cancer 74(2), 219-25

Saji J, Kurimoto N, Morita K, Nakamura M, Inoue T, Nakamura H, and Miyazawa T (2011) Comparison of 21-gauge and 22-gauge Needles for Endobronchial Ultrasound-Guided Transbronchial Needle Aspiration of Mediastinal and Hilar Lymph Nodes. Journal of Bronchology and Interventional Pulmonology 18(3), 239-246

Schreiber G, and McCrory D C (2003) Performance characteristics of different modalities for diagnosis of suspected lung cancer: Summary of published evidence. Chest 123(1 SUPPL.), 115S-128S

Soja J, Szlubowski A, Kocon P, Czajkowski W, Grzanka P, Tomaszewska R, Cmiel A, and Kuzdzal J (2010) Usefulness of transbronchial needle aspiration for initial lung cancer staging. Polskie Archiwum Medycyny Wewnetrznej 120(7-8), 264-269

Szlubowski A, Soja J, Kocon P, Talar P, Czajkowski W, Rudnicka-Sosin L, Cmiel A, and Kuzdzal J (2012) A comparison of the combined ultrasound of the mediastinum by use of a single ultrasound bronchoscope versus ultrasound bronchoscope plus ultrasound gastroscope in lung cancer staging: a prospective trial. Interactive Cardiovascular & Thoracic Surgery 15(3), 442-6; discussion 446

Trisolini R, Cancellieri A, Tinelli C, de Biase , D , Valentini I, Casadei G, Paioli D, Ferrari F, Gordini G, Patelli M, and Tallini G (2015) Randomized Trial of Endobronchial Ultrasound-Guided Transbronchial Needle Aspiration With and Without Rapid On-site Evaluation for Lung Cancer Genotyping. Chest 148(6), 1430-7

Wagner E D, Ramzy I, Greenberg S D, and Gonzalez J M (1989) Transbronchial fine-needle aspiration. Reliability and limitations. American Journal of Clinical Pathology 92(1), 36-41

Xi Y, Fan J, Che D, Zhai K, Ren T, Feng X, Shang L, Hu J, Yu Y, and Meng Q (2017) Distant metastasis and survival outcomes after computed tomography-guided needle biopsy in resected stage I-III non-small cell lung cancer. Journal of Cancer 8(16), 3356-3361

Yarmus L, Kloot T, Lechtzin N, Napier M, Dressel D, and Feller-Kopman D (2011) A randomized prospective trial of the utility of rapid on-site evaluation of transbronchial needle aspirate specimens. Journal of bronchology & interventional pulmonology 18(2), 121-127

Yarmus L, Akulian J, Ortiz R, Thompson R, Oakjones-Burgess K, Arias S, Semaan R, Feller-Kopman D, Lee H, and Wang Kp (2015) A randomized controlled trial evaluating airway inspection effectiveness during endobronchial ultrasound bronchoscopy. Journal of thoracic disease 7(10), 1825-1832

Yasuda I, Kato T, Asano F, Okubo K, Omar S, Kako N, Yasuda S, Sano K, Soehendra N, and Moriwaki H (2009) Mediastinal lymph node staging in potentially resectable non-small cell lung cancer: a prospective comparison of CT and EUS/EUS-FNA. Respiration 78(4), 423-31

Zhang R, Ying K, Shi L, Zhang L, and Zhou L (2013) Combined endobronchial and endoscopic ultrasound-guided fine needle aspiration for mediastinal lymph node staging of lung cancer: a meta-analysis. European Journal of Cancer 49(8), 1860-7

Health Economic studies – Included

Luque, M., Díez, F. J., & Disdier, C. (2016). Optimal sequence of tests for the mediastinal staging of non-small cell lung cancer. BMC medical informatics and decision making, 16(1), 9.

Navani, N., Nankivell, M., Lawrence, D. R., Lock, S., Makker, H., Baldwin, D. R., ... & Janes, S. M. (2015). Lung cancer diagnosis and staging with endobronchial ultrasound-guided transbronchial needle aspiration compared with conventional approaches: an open-label, pragmatic, randomised controlled trial. The Lancet Respiratory Medicine, 3(4), 282-289.

Sharples, L. D., Jackson, C., Wheaton, E., Griffith, G., Annema, J. T., Dooms, C., ... & Buxton, M. (2012). Clinical effectiveness and cost-effectiveness of endobronchial and endoscopic ultrasound relative to surgical staging in potentially resectable lung cancer: results from the ASTER randomised controlled trial.

NICE (2011) Clinical Guideline 121 Lung cancer: diagnosis and management. Appendix 4: Economic model to compare different testing strategies to stage the mediastinum in patients with NSCLC

Health Economic studies – Excluded

Bongers, M. L., Coupé, V. M., De Ruysscher, D., Oberije, C., Lambin, P., & Uyl-de Groot, C. A. (2015). Individualized Positron Emission Tomography–Based Isotoxic Accelerated Radiation Therapy Is Cost-Effective Compared With Conventional Radiation Therapy: A Model-Based Evaluation. International Journal of Radiation Oncology* Biology* Physics, 91(4), 857-865.

Czarnecka-Kujawa, K., Rochau, U., Siebert, U., Atenafu, E., Darling, G., Waddell, T. K., ... & Yasufuku, K. (2017). Cost-effectiveness of mediastinal lymph node staging in non–small cell lung cancer. The Journal of Thoracic and Cardiovascular Surgery, 153(6), 1567-1578.

Deppen, S. A., Davis, W. T., Green, E. A., Rickman, O., Aldrich, M. C., Fletcher, S., ... & Grogan, E. L. (2014). Cost-effectiveness of initial diagnostic strategies for pulmonary nodules presenting to thoracic surgeons. The Annals of thoracic surgery, 98(4), 1214-1222.

Dietlein, M., Weber, K., Gandjour, A., Moka, D., Theissen, P., Lauterbach, K. W., & Schicha, H. (2000). Cost-effectiveness of FDG-PET for the management of potentially operable non-small cell lung cancer: priority for a PET-based strategy after nodal-negative CT results. European journal of nuclear medicine, 27(11), 1598-1609.

Dietlein, M., Weber, K., Gandjour, A., Moka, D., Theissen, P., Lauterbach, K. W., & Schicha, H. (2000). Cost-effectiveness of FDG-PET for the management of solitary pulmonary nodules: a decision analysis based on cost reimbursement in Germany. European journal of nuclear medicine, 27(10), 1441-1456.

Esnaola, N. F., Lazarides, S. N., Mentzer, S. J., & Kuntz, K. M. (2002). Outcomes and cost-effectiveness of alternative staging strategies for non–small-cell lung cancer. Journal of clinical oncology, 20(1), 263-273.

Han, Y., Xiao, H., Zhou, Z., Yuan, M., Zeng, Y., Wu, H., & Fang, Y. (2015). Cost-effectiveness analysis of strategies introducing integrated 18F-FDG PET/CT into the mediastinal lymph node staging of non-small-cell lung cancer. Nuclear medicine communications, 36(3), 234-241.

Lejeune, C., Al Zahouri, K., Woronoff-Lemsi, M. C., Arveux, P., Bernard, A., Binquet, C., & Guillemin, F. (2005). Use of a decision analysis model to assess the medicoeconomic implications of FDG PET imaging in diagnosing a solitary pulmonary nodule. The European Journal of Health Economics, 6(3), 203-214.

Gómez León, N., Escalona, S., Bandrés, B., Belda, C., Callejo, D., & Blasco, J. A. (2014). 18f-fluorodeoxyglucose positron emission tomography/computed tomography accuracy in the staging of non-small cell lung cancer: Review and cost-effectiveness. Radiology Research and Practice, 2014.

Medford, A. R. L., Agrawal, S., Free, C. M., & Bennett, J. A. (2010). A prospective study of conventional transbronchial needle aspiration: performance and cost utility. Respiration, 79(6), 482-489.Meyers, B. F., Haddad, F., Siegel, B. A., Zoole, J. B., Battafarano, R. J., Veeramachaneni, Patterson, G. A. (2006). Cost-effectiveness of routine mediastinoscopy in computed tomography–and positron emission tomography–screened patients with stage I lung cancer. The Journal of thoracic and cardiovascular surgery, 131(4), 822-829.

Navani, N., & Janes, S. M. (2013). Endobronchial Ultrasound–guided Transbronchial Needle Aspiration for Lymphoma: The Final Frontier.

Navani, N., Nankivell, M., Woolhouse, I., Harrison, R. N., Munavvar, M., Oltmanns, Janes, S. M. (2011). Endobronchial ultrasound-guided transbronchial needle aspiration for the diagnosis of intrathoracic lymphadenopathy in patients with extrathoracic malignancy: a multicenter study. Journal of Thoracic Oncology, 6(9), 1505-1509.

Navani, Neal, et al. "Endobronchial ultrasound–guided transbronchial needle aspiration prevents mediastinoscopies in the diagnosis of isolated mediastinal lymphadenopathy: a prospective trial." American journal of respiratory and critical care medicine 186.3 (2012): 255-260.

Rintoul, Robert C., et al. "Cost effectiveness of endosonography versus surgical staging in potentially resectable lung cancer: a health economics analysis of the ASTER trial from a European perspective." Thorax (2013): thoraxjnl-2013.

Sari, A. A., Ravaghi, H., Mobinizadeh, M., & Sarvari, S. (2013). The cost-utility analysis of PET-scan in diagnosis and treatment of non-small cell lung carcinoma in Iran. Iranian Journal of Radiology, 10(2), 61.

Schreyögg, J., Weller, J., Stargardt, T., Herrmann, K., Bluemel, C., Dechow, T., ... & Buck, A. K. (2010). Cost-effectiveness of hybrid PET/CT for staging of non–small cell lung cancer. Journal of Nuclear Medicine, 51(11), 1668-1675.

Scott, W. J., Shepherd, J., & Gambhir, S. S. (1998). Cost-effectiveness of FDG-PET for staging non–small cell lung cancer: a decision analysis. The Annals of thoracic surgery, 66(6), 1876-1884.

Søgaard, R., Fischer, B. M. B., Mortensen, J., Rasmussen, T. R., & Lassen, U. (2013). The Optimality of Different Strategies for Supplemental Staging of Non–Small-Cell Lung Cancer: A Health Economic Decision Analysis. Value in health, 16(1), 57-65.

van Loon, J., Grutters, J. P., Wanders, R., Boersma, L., Dingemans, A. M. C., Bootsma, G., ... & Snoep, G. (2010). 18 FDG-PET-CT in the follow-up of non-small cell lung cancer patients after radical radiotherapy with or without chemotherapy: an economic evaluation. European Journal of Cancer, 46(1), 110-119.

Wang, Y. T., & Huang, G. (2012). Is FDG PET/CT cost-effective for pre-operation staging of potentially operative non-small cell lung cancer?–from Chinese healthcare system perspective. European journal of radiology, 81(8), e903-e909.

Appendix I – Health Economics Evidence Tables

Study, population, country and quality	Data sources	Other comments	Cost (SD)	Effect		Conclusions	Uncertainty
Navani et al. (2015) Patients who had undergone a CT scan and had suspected stage I to IIIA lung cancer. Study conducted in the UK. **Partially applicable [a, c]** **Potentially serious limitations [b, d, e]**	**Treatment effects** Taken from the LUNG-BOOST, an open-label, multicentre, pragmatic, randomised controlled trial (NCT00652769). N=133. N=66 to EBUS-TBN and n=67 to conventional diagnosis and staging (CDS, (from which one later withdrew consent). **Costs and resource use** Unit costs were obtained from NHS reference costs, NICE 2011 lung cancer guideline, and a published study; these were multiplied by the resource use and summed across all resource items. Price year 2010-2011. **Utility**	The primary endpoint was the time from first outpatient appointment with the respiratory specialist to treatment decision by the multidisciplinary team, after completion of the diagnosis and staging procedures. Analysis took a UK NHS perspective.	Conventional diagnosis and staging (n=66) 2,348 £GBP (192.20) Endobronchial ultrasound-guided transbronchial needle aspiration 2,407 £GBP (180.50) The median time to treatment decision was shorter with EBUS-TBNA (14 days; 95% CI 14–15) than with CDS (29 days; 23–35) resulting in a hazard ratio of 1·98, (1·39–2·82, p<0·0001).			"The results of the cost analysis suggested that use of EBUS-TBNA as an initial investigation after a CT scan was not more expensive than CDS. Because patients in the EBUS group of the trial had an earlier treatment decision (the primary outcome), we can conclude that EBUS-TBNA was more effective for the same cost, and was therefore cost-effective."	No sensitivity analysis was conducted.

Study, population, country and quality	Data sources	Other comments	Cost (SD)	Effect		Conclusions	Uncertainty
	Utility not measured or expressed in terms of QALYs.						

a) QALYs as per the NICE reference case were not used to measure effectiveness.
b) An incremental cost-effectiveness analysis could not be conducted in line with the NICE reference case.
c) The population was not necessarily comprised of people with an 'intermediate' probability of mediastinal malignancy as per the review protocol for this question
d) No analysis exploring uncertainty in the cost conclusions was conducted
e) No longer term cost consequences were reported

Study, population, country and quality	Data sources	Other comments	Cost (95% CI)	Effect (95% CI)		Conclusions	Uncertainty
Sharples et al. (2012) Patients requiring mediastinal staging of lung cancer. Patients had known or suspected NSCLC with suspected	**Treatment effects** Take from the ASTER, a prospective randomised controlled trial. (n=241). Surgical staging n=118. Endosonography n=123. Mean age was 64.5 years (SD 8.9). **Costs and resource use**	Analysis took a UK NHS perspective. 6-month time horizon post randomisation. Discounting not relevant.	Endosonography followed by Surgical Staging 10,808 £GBP (9,843 to 11,764)	0.348 QALYs (0.321 to 0.373)		Because of the very small QALY difference, the authors concluded that an ICER could not be estimated but 63% of bootstrapped samples showed endosonography dominated surgical	The probabilistic sensitivity analysis, showed that 63% of bootstrapped samples showed endosonography dominated (which means it was less expensive and
			Surgical Staging Alone 11,735 £GBP (10,843 to 12-647)	0.342 QALYs (0.316 to 0.367)			

Study, population, country and quality	Data sources	Other comments	Cost (95% CI)	Effect (95% CI)		Conclusions	Uncertainty
mediastinal lymph node N2 or N3 involvement. Study population from the ASTER RCT. Study conducted in the UK, The Netherlands, Belgium	Resource use was collected in terms of numbers of procedures done, (surgical, radiotherapy, chemotherapy) treatments administered, hospital and hospice stays. Costs were taken from the Department of Health (DoH) NHS reference costs 2008-2009. Estimates of endosonography was estimated by Papworth Hospital finance department. Price year 2008-2009. **Utility** Measured using the EQ-5D, in line with the NICE reference case. Utility measured at baseline, end of staging, 2 months and 6 months.	Funded by the NIHR HTA programme. Analysis also partly reported in Rintoul et al. (2013)	**Incremental cost (95% CI)**	**Incremental effect (95% CI)**	**ICER**	staging and endosonography was cost-effective at a threshold of £30,000/QALY in 99.9% of samples.	produced more benefit compared to) surgical staging and endosonography was cost-effective at a threshold of £30,000/QALY in 99.9% of samples.
			Endosonography followed by Surgical Staging vs Surgical Staging Alone				
			-927 £GBP (-2246 to 394)	0.00652 QALYs (-0.0298 to 0.0418)	Endosonography followed by Surgical Staging **Dominant**		
Directly applicable							
Potentially serious limitations [a, b, c]							

a) The costs related to combined endosonography as calculated by Papworth hospital appears to be lower than the cost of EBUS-TBNA alone as per the NICE lung cancer 2011 guidelines. The committee were unsure of the justification for this.
b) The analysis had a short time horizon so is potentially missing relevant longer term costs and QALYs
c) Complete cost and QALY information was only available for 47% of patients in each arm

Study, population, country and quality	Data sources	Other comments	Model Results	Conclusions	Uncertainty
Luque et al. (2016) Patients who require staging for suspected lung cancer. Model created for a Spanish health care setting. **Partially applicable** [b, c] **Very serious limitations** [a, d]	**Effects** Sensitivity and specificity for +ve CT scan; TBNA – Silvestri et al. (2013) PET – Gould et al. (2003) EBUS – Admas et al. (2009) EUS – Micames et al. (2007) MED – Silvestri et al. (2013) Sensitivity and specificity for -ve CT scan; TBNA – Disdier et al. (2001) PET – Gould et al. (2003) EBUS – Herth et al. (2008) MED– Silvestri et al. (2013)	This was a model based analysis, using an influence diagram (ID) that represents the possible tests, their costs, and their outcomes. This model is equivalent to a decision tree containing millions of branches. In the first evaluation, the authors only took into account the clinical outcomes (effectiveness). In the second, the authors used a willingness-to-pay of €30,000 per quality adjusted life year (QALY) to convert economic costs into effectiveness.	"Two strategies were obtained using two different criteria. When considering only effectiveness, a positive computed tomography (CT) scan must be followed by a transbronchial needle aspiration (TBNA), an endobronchial ultrasound (EBUS), and an endoscopic ultrasound (EUS). When the CT scan is negative, a positron emission tomography (PET), EBUS, and EUS are performed. If the TBNA or the PET is positive, then a mediastinoscopy is performed only if the EBUS and EUS are negative. If the TBNA or the PET is negative, then a mediastinoscopy is performed only if the EBUS and the EUS give contradictory results. When taking into account economic costs, a positive CT scan is followed by a TBNA; an EBUS is done only when the CT scan or the TBNA is negative. This recommendation of performing a TBNA in certain cases should be discussed by the pneumology community because TBNA is a cheap technique that could avoid an EBUS, an expensive test, for many patients."	"We have determined the optimal sequence of tests for the mediastinal staging of NSCLC by considering sensitivity, specificity, and the economic cost of each test. The main novelty of our study is the recommendation of performing TBNA whenever the CT scan is positive. Our model is publicly available so that different experts can populate it with their own parameters and re-examine its conclusions. It is therefore proposed	The model incorporated first order uncertainty (examined the random variability in outcomes between identical patients) and second order uncertainty (examined the uncertainty in estimation of the parameter of interest). Although the authors did not provide numerical value for the results, they concluded that the main finding of these analyses is that the resulting strategy is robust to the uncertainty of the numerical

Study, population, country and quality	Data sources	Other comments	Model Results	Conclusions	Uncertainty
	Costs and resource use Costs of tests were taken from ORDEN (2013), Gómez León (2014), Castelao Naval (2013), Kunst (2008), Navani (2009). Costs were expressed in Euros€. **Utility** Morbidities were express in QALYs. Taken from Holty (2005), Von Bartheld (2014), Silvestri (2013)			as an evidence-based instrument for reaching a consensus."	parameters because only the specificity of the EBUS when the CT scan is negative had a significant impact on the optimal strategy.

a) Costs and QALYs associated with each alternate recommended pathway are not given in the results section of the paper and sensitivity analysis are not presented in the conventional sense. It is therefore difficult to assess the face validity of the results, given the new and highly complex modelling method used in this study.
b) Costs for each of the diagnostic tests do not appear to be broadly in line with costs obtained for the UK NHS from other sources.
c) The study setting is the Spanish healthcare system, which is somewhat different from the English setting.
d) The model only has 3 treatment states, thoracotomy, chemoradiotherapy and no treatment and it is unclear whether these were appropriate and whether the costs and QALYs were taken from a relevant health system to the UK.

Study, population, country and quality	Data sources	Other comments	Model Results	Conclusions	Uncertainty
NICE Lung Cancer Guideline 2011 **Directly applicable** **Very serious limitations** [a, b, c]	Prevalence of NM stages – committee assumptions Sensitivity/Specificity of Diagnostic Tests – committee assumptions Treatment options received – NCLA registry data Overall survival – NCLA registry data Utility losses from procedures – committee assumptions Long term utility estimates – Sources from NICE TA162, TA181, TA184 Costs – EBUS micro costed, other tests from relevant UK HRG codes, treatment costs from HRGs, BNF and NICE TA181.	The economic model built for the 2011 NICE guideline examined a number of sequential testing strategies for 3 populations; those with a low, intermediate and high probability of mediastinal malignancy. Only the intermediate population is of relevance for this update.	For the intermediate population the model concludes that the most cost effective strategy is PET-CT followed by conventional TBNA, the second most cost effective strategy is neck ultrasound followed by PET-CT and conventional TBNA.	The committee noted a number of limitations with the model. Importantly, more accurate testing strategies did not lead to better outcomes for patients because false negatives were modelled to have the same outcomes as true negatives. They noted that many of the important parameters were based on assumptions but agreed it provided useful evidence in building a diagnostic pathway.	The model was robust to one way sensitivity analysis on a number of important parameters but no sensitivity analysis was conducted on the assumed diagnostic accuracy data and no probabilistic sensitivity analysis was conducted.

a) The cost differential between conventional TBNA and EBUS (£162 vs £1,365) was far larger than has been suggested by the costs analysis conducted for this guideline (see appendix J). Given that the results of the model appear highly influenced by the costs of the tests, this is an important limitation.
b) A number of crucial parameters, including the diagnostic accuracy of the tests were based on committee assumptions.
c) The modelled consequences for false negative patients may have been highly unrealistic as greater accuracy did not lead to an increase in QALYs.

Appendix J – Unit Costs of TBNA, EBUS-TBNA and EUS-FNA

Table 5: Test Costs drawn from published sources

Test	Cost	SD	Year	Source
Combined EBUS-TBNA and EUS-FNA	£ 1,237		2012	ASTER RCT (Sharples 2012) p11
EBUS-TBNA	£ 1,365		2011	NICE Lung Cancer Guideline 2011 Costing Report
EBUS-TBNA	£1,382		2012	Navani et al. 2012 (supplemental data)
Mediastinoscopy	£ 3,056	(IQR £2,360 to £3,652)	2012	ASTER RCT (Sharples 2012) p11
Thoracotomy	£ 6,525	(IQR £5,917 to £6,903)	2012	ASTER RCT (Sharples 2012) p11
TBNA	£ 423		2010	Medford et al. 2010
TBNA	£162		2011	NICE Lung Cancer Guideline 2011
TBNA	€80		2016	Luque et al. 2016

Table 6: Micro costing of EBUS-TBNA from Navani 2012 (supplementary data)

Resource	Cost per year (£)	Cost per procedure (£)	Inflated to 2017 prices	Notes
Capital costs of 2 EBUS echoendoscopes	£28,000	£112	£123	Total cost of £140,000 (including 1 processor) assumed to be spread over 5 years
EBUS-TBNA needle	£43,750	£175	£193	Source: manufacturer's price
Maintenance contract	£9,000	£36	£40	Source: UCLH
2 Consultants for 2.5 sessions per week	£50,000	£200	£220	Source: UCLH
2 Nurses, 1 health care assistant, 1 recovery nurse per session	£68,750	£275	£303	Source: UCLH

Resource	Cost per year (£)	Cost per procedure (£)	Inflated to 2017 prices	Notes
Sterilisation	£13,750	£55	£61	Source: UCLH
Pathology	£36,250	£145	£160	Source: UCLH
Administration	£10,000	£40	£44	Source: UCLH
Overheads (endoscopy suite, portering, facilities, drug costs) and Indirect costs	£86,000	£344	£379	Source: UCLH
Total cost of EBUS-TBNA	£345,500	**£1,382**	**£1,523**	

Table 7: Conventional TBNA Costs

Item	Cost	Cost per procedure (£)	Source
Cost of EBUS TBNA Needle (pack of 5 for olympus)	£1,089	£218	Source: NHS Supply Chain (Dec 2017)
Cost of conventional TBNA Needle (pack of 5)	£245	£49	Source: NHS Supply Chain (Dec 2017)
Micro-cost of a conventional TBNA		**£1,216**	Calculated = EBUS-TBNA minus per procedure costs of EBUS scope and maintenance contract and the difference in the prices of the needles
Difference between conventional TBNA and EBUS (lower estimate)		**£307**	Calculated (Navani needle price)
Difference between conventional TBNA and EBUS (higher estimate)		**£332**	Calculated (NHS Supply chain needle price)

Appendix K – Mediastinal staging of non-small cell lung cancer in patients being considered for radical treatment