PET-CT for assessing mediastinal lymph node involvement in patients with suspected resectable non-small cell lung cancer (Review)

Schmidt-Hansen M, Baldwin DR, Hasler E, Zamora J, Abraira V, Roqué i Figuls M

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2014, Issue 11

http://www.thecochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

PET-CT for assessing mediastinal lymph node involvement in patients with suspected resectable non-small cell lung cancer

Mia Schmidt-Hansen[1], David R Baldwin[2], Elise Hasler[1], Javier Zamora[3], Víctor Abraira[3], Marta Roqué i Figuls[4]

[1]National Collaborating Centre for Cancer, Cardiff, UK. [2]Department of Respiratory Medicine, Nottingham University Hospitals, NHS Trust, Nottingham City Hospital, Nottingham, UK. [3]Clinical Biostatistics Unit, Ramon y Cajal Institute for Health Research (IRYCIS), CIBER Epidemiology and Public Health (CIBERESP) and Cochrane Collaborating Centre, Madrid, Spain. [4]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau (IIB Sant Pau), CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain

Contact address: Mia Schmidt-Hansen, National Collaborating Centre for Cancer, 2nd Floor, Park House, Greyfriars Road, Cardiff, CF10 3AF, UK. Mia.Schmidt-Hansen@wales.nhs.uk.

Editorial group: Cochrane Lung Cancer Group.
Publication status and date: New, published in Issue 11, 2014.
Review content assessed as up-to-date: 30 April 2013.

Citation: Schmidt-Hansen M, Baldwin DR, Hasler E, Zamora J, Abraira V, Roqué i Figuls M. PET-CT for assessing mediastinal lymph node involvement in patients with suspected resectable non-small cell lung cancer. *Cochrane Database of Systematic Reviews* 2014, Issue 11. Art. No.: CD009519. DOI: 10.1002/14651858.CD009519.pub2.

ABSTRACT

Background

A major determinant of treatment offered to patients with non-small cell lung cancer (NSCLC) is their intrathoracic (mediastinal) nodal status. If the disease has not spread to the ipsilateral mediastinal nodes, subcarinal (N2) nodes, or both, and the patient is otherwise considered fit for surgery, resection is often the treatment of choice. Planning the optimal treatment is therefore critically dependent on accurate staging of the disease. PET-CT (positron emission tomography-computed tomography) is a non-invasive staging method of the mediastinum, which is increasingly available and used by lung cancer multidisciplinary teams. Although the non-invasive nature of PET-CT constitutes one of its major advantages, PET-CT may be suboptimal in detecting malignancy in normal-sized lymph nodes and in ruling out malignancy in patients with coexisting inflammatory or infectious diseases.

Objectives

To determine the diagnostic accuracy of integrated PET-CT for mediastinal staging of patients with suspected or confirmed NSCLC that is potentially suitable for treatment with curative intent.

Search methods

We searched the following databases up to 30 April 2013: *The Cochrane Library*, MEDLINE via OvidSP (from 1946), Embase via OvidSP (from 1974), PreMEDLINE via OvidSP, OpenGrey, ProQuest Dissertations & Theses, and the trials register www.clinicaltrials.gov. There were no language or publication status restrictions on the search. We also contacted researchers in the field, checked reference lists, and conducted citation searches (with an end-date of 9 July 2013) of relevant studies.

Selection criteria

Prospective or retrospective cross-sectional studies that assessed the diagnostic accuracy of integrated PET-CT for diagnosing N2 disease in patients with suspected resectable NSCLC. The studies must have used pathology as the reference standard and reported participants as the unit of analysis.

Data collection and analysis

Two authors independently extracted data pertaining to the study characteristics and the number of true and false positives and true and false negatives for the index test, and they independently assessed the quality of the included studies using QUADAS-2. We calculated sensitivity and specificity with 95% confidence intervals (CI) for each study and performed two main analyses based on the criteria for test positivity employed: *Activity > background* or *SUVmax ≥ 2.5* (SUVmax = maximum standardised uptake value), where we fitted a summary receiver operating characteristic (ROC) curve using a hierarchical summary ROC (HSROC) model for each subset of studies. We identified the average operating point on the SROC curve and computed the average sensitivities and specificities. We checked for heterogeneity and examined the robustness of the meta-analyses through sensitivity analyses.

Main results

We included 45 studies, and based on the criteria for PET-CT positivity, we categorised the included studies into three groups: *Activity > background* (18 studies, N = 2823, prevalence of N2 and N3 nodes = 679/2328), *SUVmax ≥ 2.5* (12 studies, N = 1656, prevalence of N2 and N3 nodes = 465/1656), and *Other/mixed* (15 studies, N = 1616, prevalence of N2 to N3 nodes = 400/1616). None of the studies reported (any) adverse events. Under-reporting generally hampered the quality assessment of the studies, and in 30/45 studies, the applicability of the study populations was of high or unclear concern.

The summary sensitivity and specificity estimates for the '*Activity > background* PET-CT positivity criterion were 77.4% (95% CI 65.3 to 86.1) and 90.1% (95% CI 85.3 to 93.5), respectively, but the accuracy estimates of these studies in ROC space showed a wide prediction region. This indicated high between-study heterogeneity and a relatively large 95% confidence region around the summary value of sensitivity and specificity, denoting a lack of precision. Sensitivity analyses suggested that the overall estimate of sensitivity was especially susceptible to selection bias; reference standard bias; clear definition of test positivity; and to a lesser extent, index test bias and commercial funding bias, with lower combined estimates of sensitivity observed for all the low 'Risk of bias' studies compared with the full analysis.

The summary sensitivity and specificity estimates for the *SUVmax ≥ 2.5* PET-CT positivity criterion were 81.3% (95% CI 70.2 to 88.9) and 79.4% (95% CI 70 to 86.5), respectively.In this group, the accuracy estimates of these studies in ROC space also showed a very wide prediction region. This indicated very high between-study heterogeneity, and there was a relatively large 95% confidence region around the summary value of sensitivity and specificity, denoting a clear lack of precision. Sensitivity analyses suggested that both overall accuracy estimates were marginally sensitive to flow and timing bias and commercial funding bias, which both lead to slightly lower estimates of sensitivity and specificity.

Heterogeneity analyses showed that the accuracy estimates were significantly influenced by country of study origin, percentage of participants with adenocarcinoma, (^{18}F)-2-fluoro-deoxy-D-glucose (FDG) dose, type of PET-CT scanner, and study size, but not by study design, consecutive recruitment, attenuation correction, year of publication, or tuberculosis incidence rate per 100,000 population.

Authors' conclusions

This review has shown that accuracy of PET-CT is insufficient to allow management based on PET-CT alone. The findings therefore support National Institute for Health and Care (formally 'clinical') Excellence (NICE) guidance on this topic, where PET-CT is used to guide clinicians in the next step: either a biopsy or where negative and nodes are small, directly to surgery. The apparent difference between the two main makes of PET-CT scanner is important and may influence the treatment decision in some circumstances. The differences in PET-CT accuracy estimates between scanner makes, NSCLC subtypes, FDG dose, and country of study origin, along with the general variability of results, suggest that all large centres should actively monitor their accuracy. This is so that they can make reliable decisions based on their own results and identify the populations in which PET-CT is of most use or potentially little value.

PLAIN LANGUAGE SUMMARY

PET-CT scanning to assess the spread of non-small cell lung cancer within the chest

In the absence of distant metastasis, treatment options for non-small cell lung cancer depend on how much the disease has spread to the different lymph nodes within the chest, that is, the stage of the disease. If the cancer has not spread beyond the nearest (N1) lymph nodes, surgery is often the treatment of choice. Other treatment options for these patients include treatment with either radiotherapy, chemotherapy, or both. Planning the optimal treatment is therefore critically dependent on accurate staging of the disease. PET-CT

scanning is a non-invasive method of establishing the spread of NSCLC within the chest and elsewhere in the body, which is increasingly available and used by lung cancer multi-disciplinary teams. Although the non-invasive nature of PET-CT constitutes one of the major advantages of the test, PET-CT may be suboptimal in detecting malignancy in normal-sized lymph nodes and in ruling out malignancy in patients with coexisting inflammatory or infectious diseases. We examined the accuracy of PET-CT scanning in establishing the spread of cancer in patients with suspected or confirmed NSCLC that is potentially suitable for surgical treatment with curative intent.

We included 45 studies, and based on the criteria for a positive PET-CT scan, we performed two main analyses. In the 18 studies (2823 participants) in the *Activity > background* group, PET-CT was found to accurately identify 77.4% (95% CI 65.3 to 86.1) of the participants with NSCLC spread beyond the N1 nodes and 90.1% (95% CI 85.3 to 93.5) of the participants without spread beyond the N1 nodes. In the 12 studies (1656 participants) in the *SUVmax of ≥ 2.5* group, PET-CT accurately identified 81.3% (95% CI 70.2 to 88.9) of the participants with spread beyond the N1 nodes and 79.4% (95% CI 70 to 86.5) of the participants without spread beyond the N1 nodes. However, the results varied a lot between the studies in each analysis, and the quality and size of the studies themselves, country of study origin, percentage of participants with adenocarcinoma, FDG dose, and type of PET-CT scanner influenced the results. We believe that the results of this review show that the accuracy of PET-CT is insufficient to allow management based on PET-CT alone.

BACKGROUND

Accurately determining the diagnosis and stage of lung cancer is important to ensure that patients are offered the best possible treatment. However, the process is often complex. The symptoms and signs of lung cancer can be difficult to distinguish from those of other diseases (some of which may coexist in lung cancer patients), and many lung cancers are diagnosed via other routes (e.g., emergency or Accident & Emergency admissions; through other specialities; or as incidental findings on imaging, such as chest radiographs and computed tomography (CT)) (Department of Health 2011). The diagnosis is made by means of a variety of different biopsies and imaging techniques, some of which yield information about both diagnosis and staging (NICE 2011). The need to consider the location of the primary tumour; patient preferences; and the fitness of the patient, which itself may influence both diagnostic and treatment decisions and may require a change to the diagnostic and staging pathway, augments the complexity.

Target condition being diagnosed

A major determinant of treatment offered to patients with non-small cell lung cancer (NSCLC) is the intrathoracic (mediastinal) nodal status (for a glossary, see Appendix 1). If the disease has not spread to either the ipsilateral mediastinal nodes, subcarinal (N2) nodes, or both, and the patient is otherwise considered fit for surgery, resection is often the treatment of choice (Manser 2005). Other treatment options for these patients include combination or single-modality treatment with either radiotherapy, chemotherapy, or both (O'Rourke 2010). Planning the optimal treatment is therefore critically dependent on accurate staging of the disease.

Lung cancer staging is performed using an arsenal of different complementary tests; some of these are non-invasive (e.g., various types of imaging) (NICE 2011; Silvestri 2013), and some are invasive (e.g., surgical staging, mediastinoscopy) or minimally invasive (e.g., endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA), endoscopic ultrasound-guided fine needle aspiration (EUS-FNA), and various other methods for obtaining biopsies) (Detterbeck 2007; NICE 2011). The most definitive test is surgical staging - by means of resection of the primary tumour and systematic nodal dissection, with prior mediastinoscopy to assess the contralateral nodes. However, surgical staging is highly invasive and thus not appropriate for many patients without first acquiring further information about the likely suitability for resection with curative intent. Imaging tests (including combined positron emission tomography and CT (PET-CT)) - to assess the probability of malignant involvement and detect extrathoracic metastases and mediastinal lymph node metastasis that would preclude treatment with curative intent - will most often determine suitability for resection with curative intent. One or more biopsies may have to follow imaging findings to pathologically confirm the results of these tests. Occasionally, when imaging tests are unequivocally positive for cancer, the findings alone will be enough to exclude patients from radical treatment. This, in effect, means that only the patients who receive resection will receive the ultimate reference standard (i.e., surgical staging). Those patients who are found to have unresectable NSCLC will usually have had their cancer stage pathologically confirmed by a number of other tests that are considered suitable for the location of the affected lymph node(s).

Therefore, the reference standard for this review necessarily had to consist of a number of invasive tests that all yield pathologically

confirmable information and can be collectively considered as tests that provide cytohistological confirmation of tumour extent. The secondary aims of this review reflect our consideration of this issue as we have tried to consider potential differences in the reference standard as a source of heterogeneity between the studies.

Index test(s)

PET-CT is a non-invasive staging method of the mediastinum, which is increasingly available and used by lung cancer multidisciplinary teams. PET-CT is most commonly performed using

(^{18}F)-2-fluoro-deoxy-D-glucose (FDG) as a tracer to provide a measure of glucose uptake, with simultaneous CT to aid localisation. Before receiving a PET-CT scan, most patients will already have received a CT scan, and PET-CT is most commonly used to confirm early-stage disease in patients who have no significant nodes ($\geq$ 1 cm on the short axis) on CT or to clarify nodal status, in which case PET-CT is not always the first test after CT. That is, currently, the role of PET-CT is primarily in triaging patients, by identifying patients with no spread to the mediastinum who may therefore be candidates for resection, and distinguishing from those patients with either distant or mediastinal metastases, or both, that may need to be biopsied before their treatment plan can be developed.

Clinical pathway

NSCLC patients present with a variety of symptoms and signs as described in the NICE guidelines on referral for suspected cancer (NICE 2005). In England, about 38% of patients first present through the emergency route (i.e., Accident & Emergency or medical admissions). General practitioners urgently refer the majority of the remainder. In most cases, with the exception of those who are too ill to be helped by further diagnostic attempts, the first diagnostic step when lung cancer is suspected is imaging that is either chest radiography or multidetector CT. The latter should ideally be done with the administration of intravenous contrast with contiguous slices from the lower neck to upper abdomen. The secondary care pathway begins with this CT and a clinical assessment where the history is taken, a physical examination, and basic blood tests and lung function obtained. From this information, the first estimate of fitness is made. As part of the two-way communication with either the patient or carer (or both) the potential diagnosis is explained, and some idea about the patient's preference is formulated. The next step in the pathway is to choose the test that gives the most diagnostic and staging information with least risk of harm, provided that the patient is agreeable to this and that further information will likely help the patient. This choice is heavily dependent on what is shown by the CT, and NICE clinical guideline 121 gives detailed guidance on the most appropriate choice of test (NICE 2011; see also De Leyn 2014). The most relevant part of this guidance for the purposes of this review concerns the staging of the mediastinum, which has a separate and more detailed algorithm within the NICE guideline. Once diagnosis, stage, and fitness assessment is confirmed, treatment may be offered on the basis of this information, and follow-up is usually supervised in secondary care in liaison with community services. On relapse, patients may be reassessed, which may be as detailed as the initial work-up, but is usually less so, with treatment offered again on the basis of the findings of the reassessment. Patients in the UK have access to the support of lung cancer nurse specialists throughout the pathway, and there should be holistic needs assessment at all stages of diagnosis and treatment.

Role of index test(s)

PET-CT is central to the assessment of patients who might potentially be suitable for treatment with curative intent. This test is able to define more clearly whether lung cancer has spread to lymph nodes or further. It is in routine use in the UK and a standard of care. NICE recommends PET-CT when the CT does not show significantly enlarged lymph nodes or where nodes of intermediate probability of malignancy are seen (NICE 2011). In reality, many PET-CTs are done for larger high-probability (of cancer) nodes (on CT) prior to minimally invasive sampling, although this practice is unlikely to be cost-effective. Thus, PET-CT forms part of a sequence of tests in the work-up of patients potentially suitable for surgery and is increasingly being done early in the pathway after a baseline CT scan. However, PET-CT is not a perfect test, and it is important to quantify its accuracy and be aware of factors that might alter this.

Alternative test(s)

Other imaging modalities can provide similar information to PET-CT, and these include contrast-enhanced magnetic resonance imaging (MRI) and single photon emission-computed tomography (SPECT). Neither of these tests are as widely available as PET-CT, and importantly, unlike PET-CT, they are also not embedded in the lung cancer pathway. Other tests include the minimally invasive lymph node sampling procedures (i.e., EBUS-TBNA, EUS-FNA) and may be used ahead of PET-CT when treatment might be determined by the result from a single nodal station. Where this sample is negative for malignancy, this approach risks the need for further sampling if PET-CT suggests malignancy in other node stations. For the purposes of this review, the minimally invasive sampling techniques are not considered as alternative tests per se, but rather as part of the techniques that all provide pathological information and thereby collectively constitute the reference standard (see also Target condition being diagnosed).

Rationale

Although the non-invasive nature of PET-CT constitutes one of the major advantages of the test, PET-CT may be suboptimal in detecting malignancy in normal-sized lymph nodes and in ruling out malignancy in patients with coexisting inflammatory or infectious diseases (Cerfolio 2005; Kim 2006; Lee 2007a; Shim 2005; Tournoy 2007; Yi 2007). The role of PET-CT in the accurate staging pathway for patients with lung cancer is therefore still debated, and a crucial question is when a biopsy sample is needed to increase the sensitivity and specificity of PET-CT. Multidisciplinary teams must have a clear idea of the likelihood of false positive and negative PET-CT results in a given circumstance (in particular, the size of mediastinal nodes) in order to best manage patients and advise them whether or not a biopsy is necessary. A false negative rate that is consistently above 20% would cause clinicians to question the utility of the test. However, the question is complex. For example, in the case of detection of distant metastases in patients otherwise fit for surgery, a 20% false negative rate might lead to only one patient in 100 having futile surgery (as the baseline rate of distant metastases is around 5%). In the case of assessing mediastinal nodes by PET-CT, the overall impact will again depend on the prevalence. However, it is also noted that resection of these nodes may not necessarily mean that an operation was the wrong thing to do, as we know from the National Lung Cancer Audit that outcomes are better when patients have had surgery, even for N2 disease (NICE 2011). On balance, we have focused on nodal metastases in this review. False negative outcomes of PET-CT should only apply to nodes that are not significantly enlarged on (a prior) CT, as enlarged nodes should be biopsied. False positives are of a lesser concern since they should always be followed by a further test to confirm.

This review represents an extension to a previous review we have undertaken in this area for the 2011 NICE updated guideline on the diagnosis and treatment of lung cancer (NICE 2011); this included fewer studies and no meta-analysis.

OBJECTIVES

To determine the diagnostic accuracy of integrated PET-CT for mediastinal staging of patients with suspected or confirmed NSCLC that is potentially suitable for treatment with curative intent.

Secondary objectives

To assess potential sources of heterogeneity, including study design (e.g., retrospective/prospective, consecutive/random series); patient populations (number and characteristics, e.g., T- and N-stage, significant nodes on prior CT, country); different cut-off values for test positivity (malignancy); differences in either PET-CT image acquisition, scanning equipment, or both; and potential differences in reference standard (mediastinoscopy/pathological or surgical staging).

METHODS

Criteria for considering studies for this review

Types of studies

Prospective or retrospective cross-sectional studies that assessed the diagnostic accuracy of integrated PET-CT for diagnosing N2 disease in patients with suspected resectable NSCLC. The studies must have used pathology as the reference standard and reported participants as the unit of analysis.

Participants

Patients with suspected/confirmed NSCLC who were considered potentially suitable for primary resection. This review did not consider patients who were being restaged after induction or neoadjuvant chemotherapy.

Index tests

PET-CT carried out on the various available integrated PET-CT scanners with cut-off values for test positivity as reported in the included studies. The type of integrated PET-CT scanner, scanner manufacturer, and cut-off values did not influence whether we included a study or not; rather, as part of the secondary objectives, we examined the potential contribution of these factors to systematic between-study variation as potential sources of heterogeneity. However, we did not consider studies that employed tracers other than FDG or other nuclear medicine imaging, such as single photon emission-computed tomography (SPECT) or stand-alone PET.

Target conditions

Resectability of lung cancer depends on the locoregional spread of the disease. NSCLC is generally not considered resectable if it has spread beyond N1 disease. Thus, the target condition of this review was resectable NSCLC, which for the present purposes, was defined as NSCLC that has not spread to either the ipsilateral mediastinal lymph nodes, the subcarinal (N2) lymph nodes, or both.

Reference standards

Pathological confirmation of PET-CT results from samples obtained via either surgical resection with mediastinal sampling, mediastinoscopy, video-assisted thoracic surgery (VATS), endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA), EUS-FNA, TBNA (transbronchial needle aspiration), TTNA (transthoracic needle aspiration), biopsies of extrathoracic sites, or a combination of the aforementioned.

Search methods for identification of studies

Electronic searches

We searched the following databases up to 30 April 2013, using the search terms and strategies identified in Appendix 2:

- *The Cochrane Library* (specifically, the Cochrane Central Register of Controlled Trials (CENTRAL), the Health Technology Assessment (HTA) database, the Database of Abstracts of Reviews of Effects (DARE), and the Cochrane Methodology Register (CMR));
- MEDLINE via OvidSP (from 1946);
- Embase via OvidSP (from 1974);
- PreMEDLINE via OvidSP;
- OpenGrey; and
- ProQuest Dissertations & Theses.

We also searched the trials register http://clinicaltrials.gov for research projects in process on 30 April 2013. We used Web of Science (or Scopus if the citation was not on Web of Science) to track records citing those studies, which we included in the final review, with an end-date of 9 July 2013. There were no language or publication status restrictions on the search.

Searching other resources

We handsearched the reference lists of the included articles along with the reference lists of any relevant review articles identified through the search. We also contacted the authors of the included studies and other experts in the field of lung cancer staging for information about any ongoing or unpublished studies. We imposed no language or publication status restrictions on the search.

Data collection and analysis

Selection of studies

Firstly, one of the review authors (MSH) assessed for potential inclusion the titles and abstracts of all the studies identified by the search. This first stage of screening excluded all records that were not studies of PET-CT in patients with NSCLC. Secondly, two of the review authors (MSH and DRB) assessed for potential inclusion the titles and abstracts of the remaining records. Thirdly, two of the review authors (MSH and DRB) independently considered the full records of all potentially relevant studies for inclusion by applying the selection criteria outlined in the Types of studies section. We resolved any disagreements by discussion.

Data extraction and management

Using a standardised data extraction form, two authors (MSH and DRB or MRF) extracted data pertaining to study design, participant detail, index and reference tests, and funding (see Table 1). We resolved any disagreements by discussion. With studies where only a subgroup of the participants met the inclusion criteria for the current review, we only extracted data on this subgroup. For the comparison of the index test with the reference standard, we extracted the number of true and false positives and true and false negatives for the index test when these numbers were presented in the studies. Otherwise, we reconstructed the two-by-two table of true and false positives and negatives from the information reported in the studies, and if this was not possible, we contacted the study authors for the data.

Assessment of methodological quality

Two of three of the authors (MSH and DRB or MFR) independently assessed the quality of each study using a modified version of the QUADAS-2 tool (Whiting 2011), as outlined in Table 2. QUADAS-2 consists of four domains that each require a 'Risk of bias' judgement of low, high, or unclear. For three of these domains, a further judgement needs to be made rating concerns of applicability as low, high, or unclear in terms of how applicable the individual study results are to the question posed by the review. Signalling questions that require a yes, no, or unclear response support the 'Risk of bias' judgements. We included two additional signalling questions on our checklist:

1. Was there a clear definition of a positive result? (We included this under the 'Index test' domain.)
2. Was the study free of commercial funding?

We included the item pertaining to the definition of positive results to take into account the subjective nature of PET-CT image interpretation, which may be based on a variety of different criteria, such as extensive clinical experience, different standard uptake values (SUV), different morphological features, or a combination of the aforementioned. We included the second additional item in order to record any potential bias resulting from commercial interest in the results. We resolved any disagreements between the risk of bias and applicability concern ratings through discussion.

Statistical analysis and data synthesis

We extracted the numbers of true positives, false positives, true negatives, and false negatives for each study based only on the

ability of PET-CT to distinguish between N0 and N1 mediastinal disease and N2 and N3 mediastinal disease. Therefore, we considered both N0 and N1 disease as negatives and both N2 and N3 as positives. If PET-CT indicated N1 disease that was shown by the reference standard to be N0 disease (and vice versa), the PET-CT results were still considered a true negative because N0 and N1 disease were both considered resectable disease. The same principle applied to N2 and N3 disease, that is, if PET-CT indicated N2 disease that was shown to be N3 disease by the reference standard (and vice versa), the PET-CT results were still considered a true positive. However, if PET-CT indicated N0 or N1 disease that the reference standard showed to be N2 or N3 disease, the FDG PET result was considered a false negative. Similarly, if PET-CT indicated N2 or N3 disease that was shown by the reference standard to be N0 or N1 disease, the PET-CT result was considered to be a false positive. If data for more than one positivity threshold were reported, we extracted all the data, but only analysed the threshold most commonly used by all the studies. We only extracted data with participant as the unit of analysis, not, for example, lymph node.

We calculated sensitivity and specificity with 95% confidence intervals (CI) for each study. We plotted the estimates of the observed sensitivities and specificities together with their 95% CI in forest plots and in a receiver operating characteristic (ROC) plot of sensitivity versus 1-specificity in order to visually assess the between-study variability. We fitted a summary ROC curve using the HSROC model for the subset of studies sharing the same positivity threshold (Harbord 2007; Rutter 2001). We selected one threshold per study in the special case of a single study reporting data for more than one threshold. If the studies showed sufficient clinical homogeneity (see Investigations of heterogeneity), we derived summary accuracy estimates for the studies using the same criteria for test positivity for all participants (i.e., *SUVmax* $\geq$ *2.5*, *Activity* > *background*). In the case of different thresholds used in the studies for the analyses, we selected the most frequently used, clinically relevant threshold among the included studies. We identified the average operating point on the SROC curve and computed average sensitivities and specificities. We plotted averaged accuracy estimates with their 95% confidence ellipse and prediction region in ROC space. We had planned to compute the positive and negative likelihood ratios from the pooled estimates of sensitivity and specificity, but given the high degree of heterogeneity we found, the accuracy estimates should be interpreted with caution. As a consequence, we did not compute the likelihood ratios in order to separate the results of our review from their use in clinical practice for a specific patient (i.e., updating post-test probability after a test result).

Investigations of heterogeneity

Several factors can contribute to heterogeneity in diagnostic accuracy of a test across studies. We checked for heterogeneity as part of the planned meta-analysis. Anticipated sources of heterogeneity included study design (e.g., retrospective/prospective, consecutive/random series); FDG dose; patient populations (year, country, sample size, percentage of adenocarcinoma, country tuberculosis rate); and differences in PET-CT image acquisition or scanning equipment (or both).

We could not explore potential differences in reference standard (mediastinoscopy/pathological or surgical staging), one of the planned sources of heterogeneity, because of lack of adequate data.

We replaced another planned source of heterogeneity (different cut-off values for test positivity) by the type of test positivity (surrounding activity, SUVmax, and other criteria).

We conducted a subgroup analysis for each factor anticipated to be a heterogeneity source by including the factor as a covariate in the bivariate model (Reitsma 2005). We performed comparison of diagnostic accuracy between subgroups by testing whether either sensitivity or specificity, or both, differed in subgroups of studies defined according to the covariate. The analysis aimed to estimate valid measures of diagnostic accuracy taking into account the effect of any confounding variables. We used the non-linear mixed models (NLMIXED) (Macaskill 2004) procedure in SAS version 9.1 for Windows (SAS Institute Inc, Cary, NC, USA) to fit the HSROC and bivariate models.

Sensitivity analyses

We examined the robustness of the meta-analyses by conducting sensitivity analyses using different components of the 'Risk of bias' assessment. We performed these analyses by limiting inclusion in the meta-analysis to those studies in the primary analyses that had low risk of bias and low concerns about potential applicability. We also excluded from the analyses studies according to other characteristics that could potentially introduce bias into the results (i.e., whether a clear definition for test positivity was used and whether commercial funding was provided).

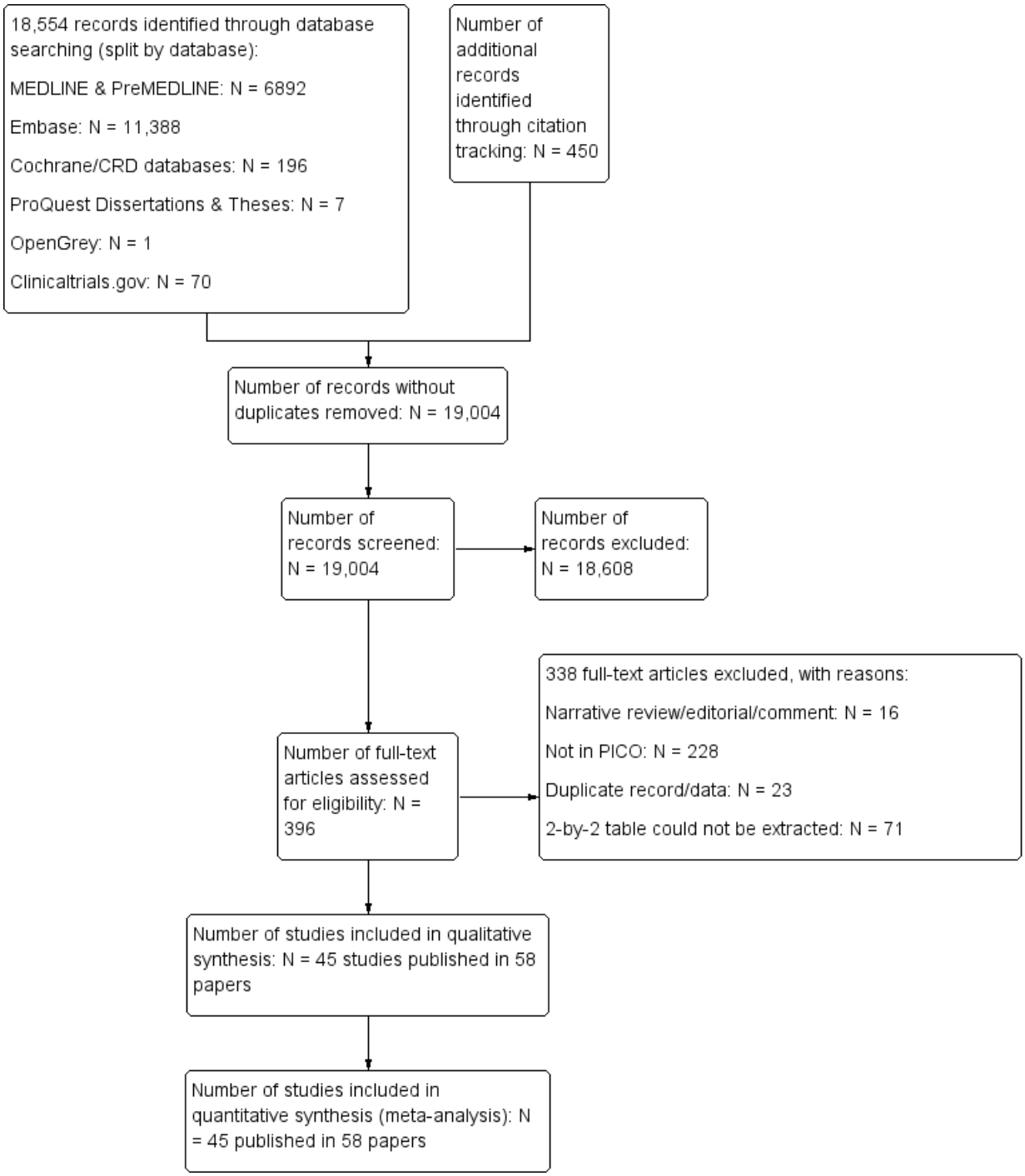
RESULTS

Results of the search

Our search strategy identified 19,004 records, of which we excluded 18,608 as not relevant, based on the title/abstract, and obtained the full publications of 396 records. Of the full publications, 45 studies published in 58 papers met the inclusion criteria while we excluded 338 articles for the following reasons: narrative review/editorial/comment (N = 16); not meeting the PICO (population, intervention, comparison, outcome) criteria (i.e., either the population or tests did not match the current target population or index/reference tests, or the study did not examine the accuracy of PET-CT for mediastinal staging) (N = 228); duplicate

record or data (N = 23); or the two-by-two table could not be extracted for patient level N0 and N1 versus N2 and N3 data (N = 71) (see also Figure 1 and the 'Characteristics of excluded studies' tables). The 45 included studies had a total of 6095 participants available for analysis (median = 112, interquartile range (IQR) = 54 to 169), 4551 of whom were N0 and N1 and 1544 participants of whom were N2 and N3. The prevalence of positive nodes (N2 and N3) varied amongst the studies, ranging from as low as 4% (Lee 2012) to 83% (Uskul 2009), with a median of 22% (IQR = 18 to 30). Thirty-two studies reported the percentage of participants with adenocarcinoma, which ranged from 20.5% to 87.2%. The studies were categorised according to the incidence rate of tuberculosis (TB, which also included HIV (human immunodeficiency virus)) as reported by the World Health Organization (WHO) (www.who.int/tb/country/data/profiles/en/index.html). Two thirds of the studies (N = 30) had incidence rates lower than 50 per 100,000 population. Half of the studies were performed in Asia (N = 22), while Europe provided 11 studies; North America, a further three studies; and nine studies were from other countries (Turkey, Egypt). All the studies were published after 2005 (2006 to 2009: N = 17; 2010 to 2011: N = 17; and 2012 to 2013: N = 11).

Figure 1. Study flow diagram



The studies also varied in which PET-CT scanner they used, with 19 studies using a Discovery scanner, 14 studies using a Biograph scanner, and the remaining 12 studies employing other/mixed/not reported scanners. We also observed between-study variation in terms of the FDG dose used for the PET-CT scans. Where the total dose was not reported directly, the data were converted to total FDG dose in MBq (megabecquerel) in the following manner: When the dose was reported as MBq/kg, we calculated a total dose for a participant weighing 70 kg. When the FDG dose was reported as a range, we used the mean value. According to these calculations, 12 studies used up to 300 MBq, 25 studies used 301 to 500 MBq, and four studies used > 500 MBq. The remaining four studies did not report FDG dose. There was little difference in injection-to-scan time between the studies (> 45 minutes: N = 1; 30 to 60 minutes: N = 1; 40 to 60 minutes: N = 1; 45 minutes: N = 2; 45 to 60 minutes: N = 1; 50 minutes: N = 3; 60 minutes: N = 26; 55 to 65 minutes: N = 1; 50 to 70 minutes: N = 2; 60 to 120 minutes: N = 1; 75 minutes: N = 1; not reported: N = 5). Twenty-nine studies used attenuation correction; one study did not; and 15 studies did not report whether they undertook attenuation correction. The included studies used different criteria for test positivity. Based on these criteria, we categorised the included studies into three groups of criteria for test positivity: *Activity > background* (18 studies, N = 2823, prevalence of N2 and N3 nodes = 679/2823), *SUVmax ≥ 2.5* (12 studies, N = 1656, prevalence of N2 and N3 nodes = 465/1656), and *Other/mixed* (15 studies, N = 1616, prevalence of N2 and N3 nodes = 400/1616). None of the studies reported (any) adverse events. For full detail study details, see the 'Characteristics of included studies' tables.

Methodological quality of included studies

We have summarised below the methodological quality of the included studies as assessed by QUADAS-2 and per study and per QUADAS-2 item in Figure 2 and Figure 3, respectively. Inspection of Figure 2 and Figure 3 reveals that a substantial amount of under-reporting in the original studies, which led to many judgements of unclear, hampered the quality of the data.

Figure 2. 'Risk of bias' and applicability concerns summary: review authors' judgements about each domain for each included study

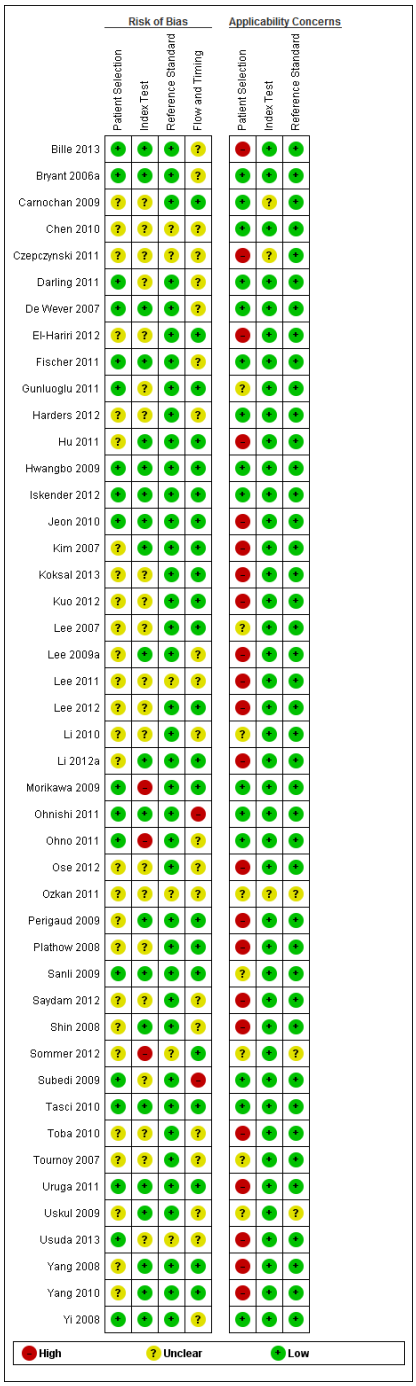
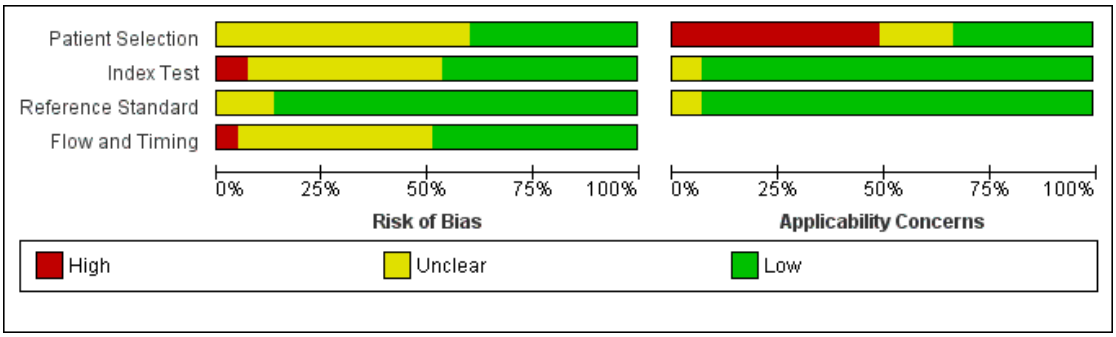


Figure 3. 'Risk of bias' and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies



Participant selection

Risk of bias

We judged participant selection to be at low risk of bias in 18 of the studies and at unclear risk of bias in the remaining 27 studies.

Applicability concerns

A substantial number of the included studies only included participants who had received resection for NSCLC (Bille 2013; Czepczynski 2011; El-Hariri 2012; Hu 2011; Jeon 2010; Kim 2007; Koksal 2013; Kuo 2012; Lee 2009a; Lee 2011; Li 2012a; Ose 2012; Perigaud 2009; Plathow 2008; Toba 2010; Uruga 2011; Usuda 2013; Yang 2008; Yang 2010), while other studies only included participants with T1 NSCLC (Lee 2012; Shin 2008) or who were retired coal workers (Saydam 2012). All of these inclusion restrictions artificially narrow the range of patients who would receive FDG/PET-CT in standard practice, in particular, the patients with N2 and N3 disease, which in turn gives rise to high concern about the applicability of the populations to the objective of this review. Eight studies did not provide enough information for this item to be rated (i.e., we classified these studies as unclear concerns about applicability), while the populations of the remaining studies were directly applicable to the current question (thus, we classified them as low concern about applicability).

Index test

Risk of bias

The index test was of low or unclear risk in the vast majority of the included studies. However, in three of the included studies, the risk of bias for the index test was high because the results were based on a posthoc specification of the optimal threshold (Morikawa 2009; Ohno 2011) or on more data than just the PET-CT images (Sommer 2012). This was along with a flexible/non-systematic use of SUVs without a general cut-off value (Sommer 2012).

Applicability concerns

We rated three of the included studies as unclear for applicability of the index test because not enough information was reported to assess this question (Carnochan 2009; Czepczynski 2011; Ozkan 2011). We considered the index test as employed by the remaining studies to be applicable to the aims of this review.

Reference standard

Risk of bias

We considered all of the included studies to be at low risk of bias with the exception of Chen 2010; Czepczynski 2011; Lee 2011; Ozkan 2011; Sommer 2012; Usuda 2013, which were all of unclear risk of bias for the reference standard.

Applicability concerns

We considered the reference standard to be applicable to the review in all the included studies apart from three (Ozkan 2011; Sommer

2012; Uskul 2009), which we rated as unclear because of a lack of information reported in the papers, making it impossible to assess the applicability of the reference standard in these cases.

Flow and timing

Risk of bias

Most of the studies were of low or unclear risk of bias, but we considered two of the studies to be at high risk of bias for flow and timing because of missing data (Ohnishi 2011; Subedi 2009).

Other assessed 'Risk of bias' items

Prespecified cut-off values for PET-CT positivity

We selected this item for preplanned sensitivity analyses to assess if the results were sensitive to whether the cut-off values for test positivity were specified a priori or posthoc. However, on appraising the included studies, it became apparent that this item did not apply to at least half of the included studies, that is, the studies that did not use an explicitly quantitative test measure (i.e., SUV). Because when no quantitative criterion has been employed, the answer to this item is 'no' without this in itself giving rise to a problem. We therefore decided to just incorporate this potential source of bias into the 'Risk of bias' assessment for the index test and to limit the assessment of the influence of this item to the sensitivity analysis of the risk of bias for the index test.

PET-CT test positivity clearly defined

Only in nine studies were the criteria for PET-CT positivity either unclearly defined (Ohno 2011; Ozkan 2011; Plathow 2008; Sommer 2012; Tournoy 2007) or not defined (Carnochan 2009; Chen 2010; Czepczynski 2011; Darling 2011); whereas, the remaining 36 studies clearly defined test positivity.

Commercial funding of the studies

Only 19 studies reported any details about funding, and of those studies, 14 had received non-commercial funding (Darling 2011; Fischer 2011; Hu 2011; Hwangbo 2009; Kuo 2012; Lee 2012; Li 2010; Li 2012a; Morikawa 2009; Shin 2008; Usuda 2013; Yang 2008; Yang 2010; Yi 2008); two had received commercial funding (Ohno 2011; Sommer 2012); and three studies reported that they had received no funding (Saydam 2012; Tournoy 2007; Uruga 2011).

Findings

Accuracy of integrated PET-CT for mediastinal staging

Figure 4, Figure 5, and Figure 6 show forest plots of PET-CT sensitivity and specificity for assessing mediastinal lymph node involvement for all the 45 studies included in the review, grouped by the criteria for test positivity employed, i.e., *Activity > background*, *SUVmax ≥ 2.5*, or *Other/mixed/unclear*. Both sensitivity and specificity estimates varied greatly within all three groups. Indeed, sensitivity estimates varied by more than 50% in all three groups, with specificity estimates varying by at least 27% within the groups.

Figure 4. Forest plot of studies with Activity > background as the criterion for test positivity

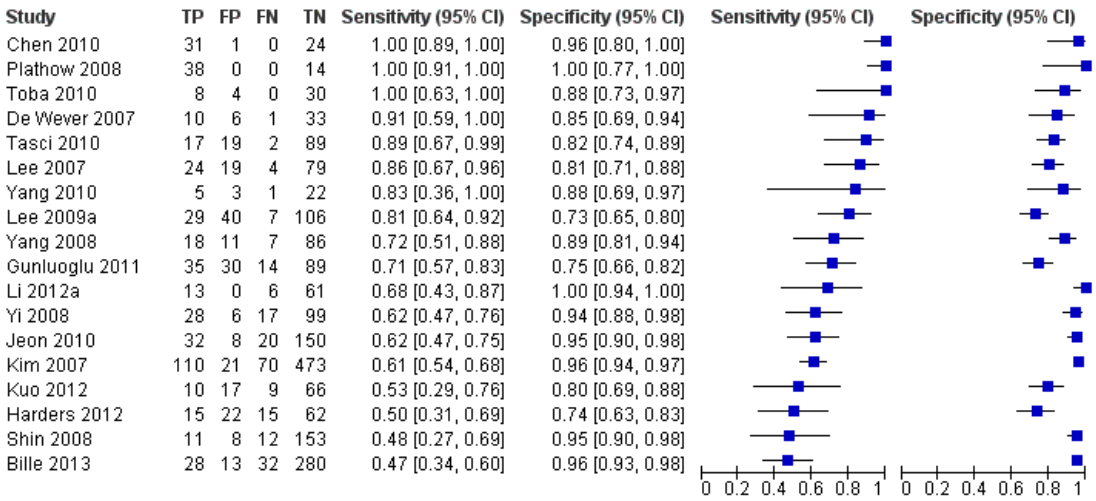


Figure 5. Forest plot of studies with SUVmax $\geq$ 2.5 as the criterion for test positivity

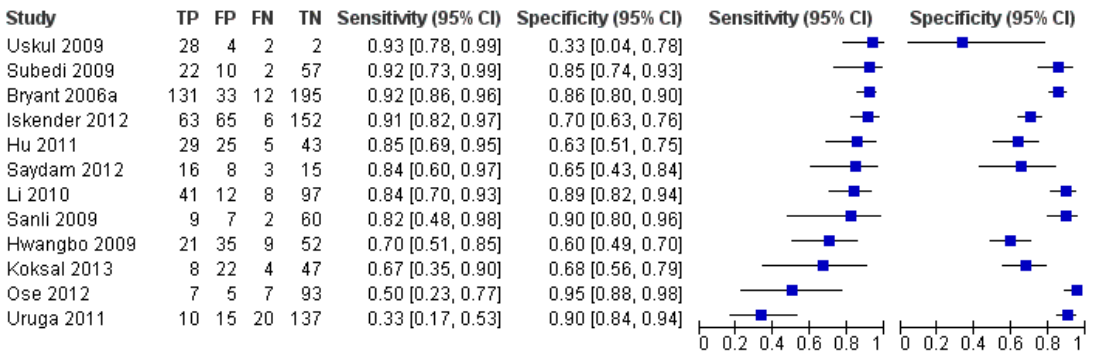
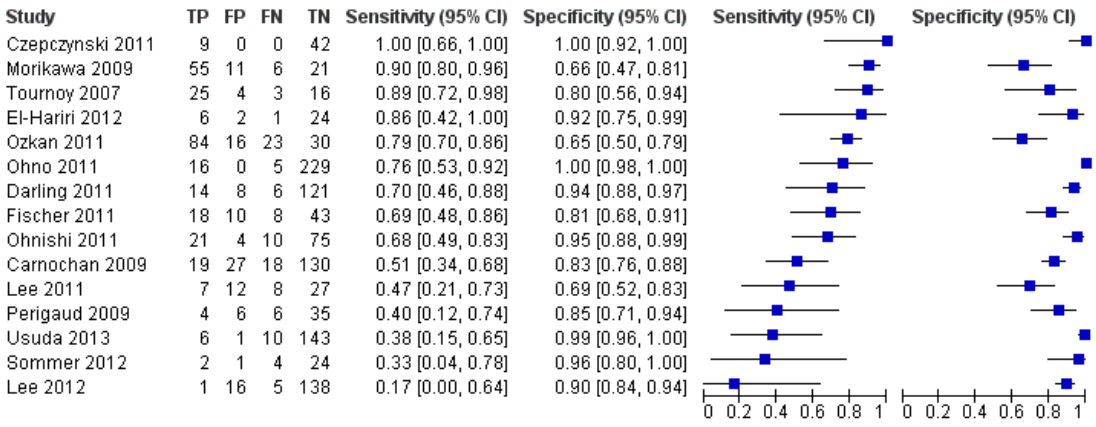
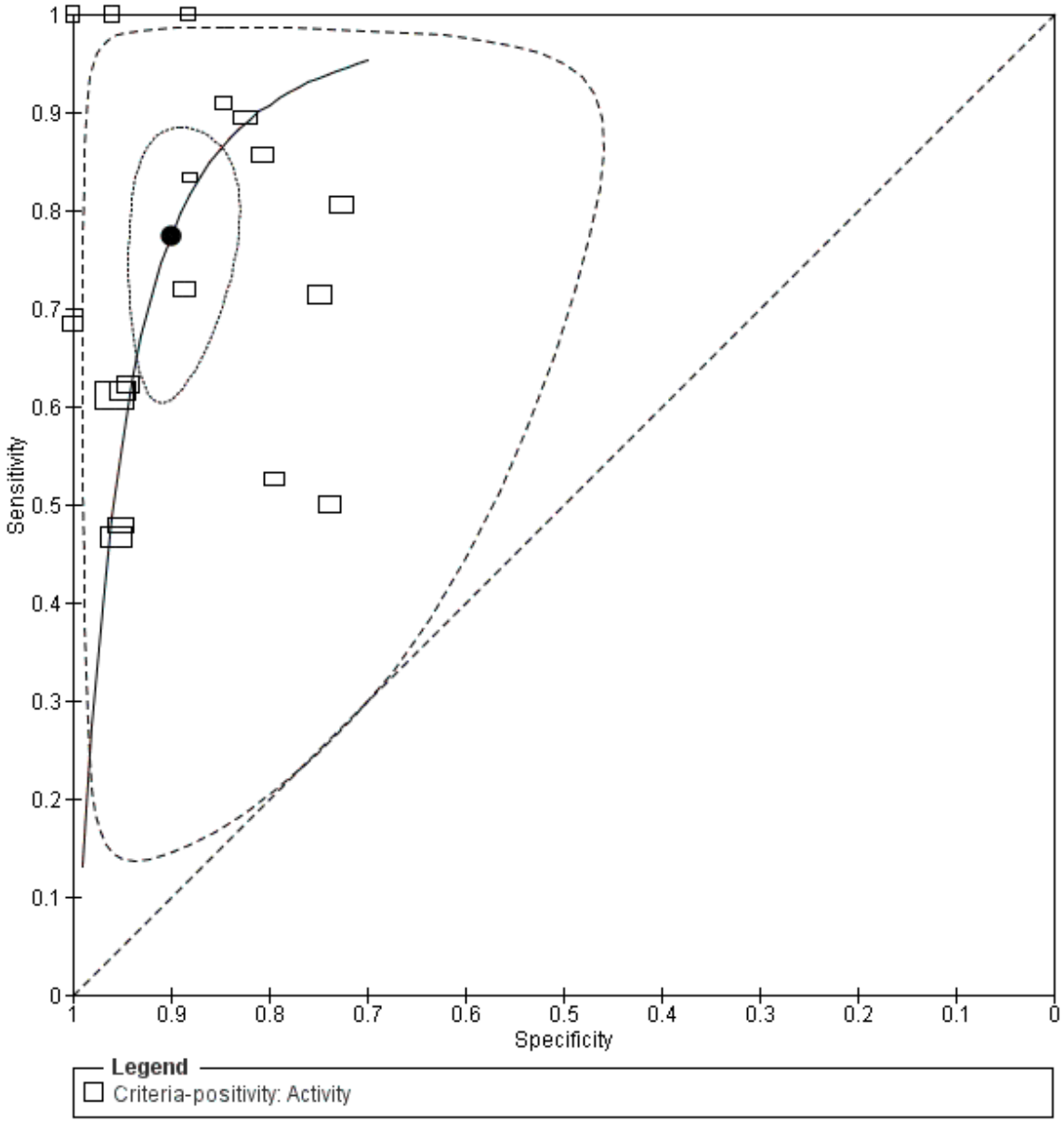


Figure 6. Forest plot of studies with Other/mixed/unclear criteria for test positivity



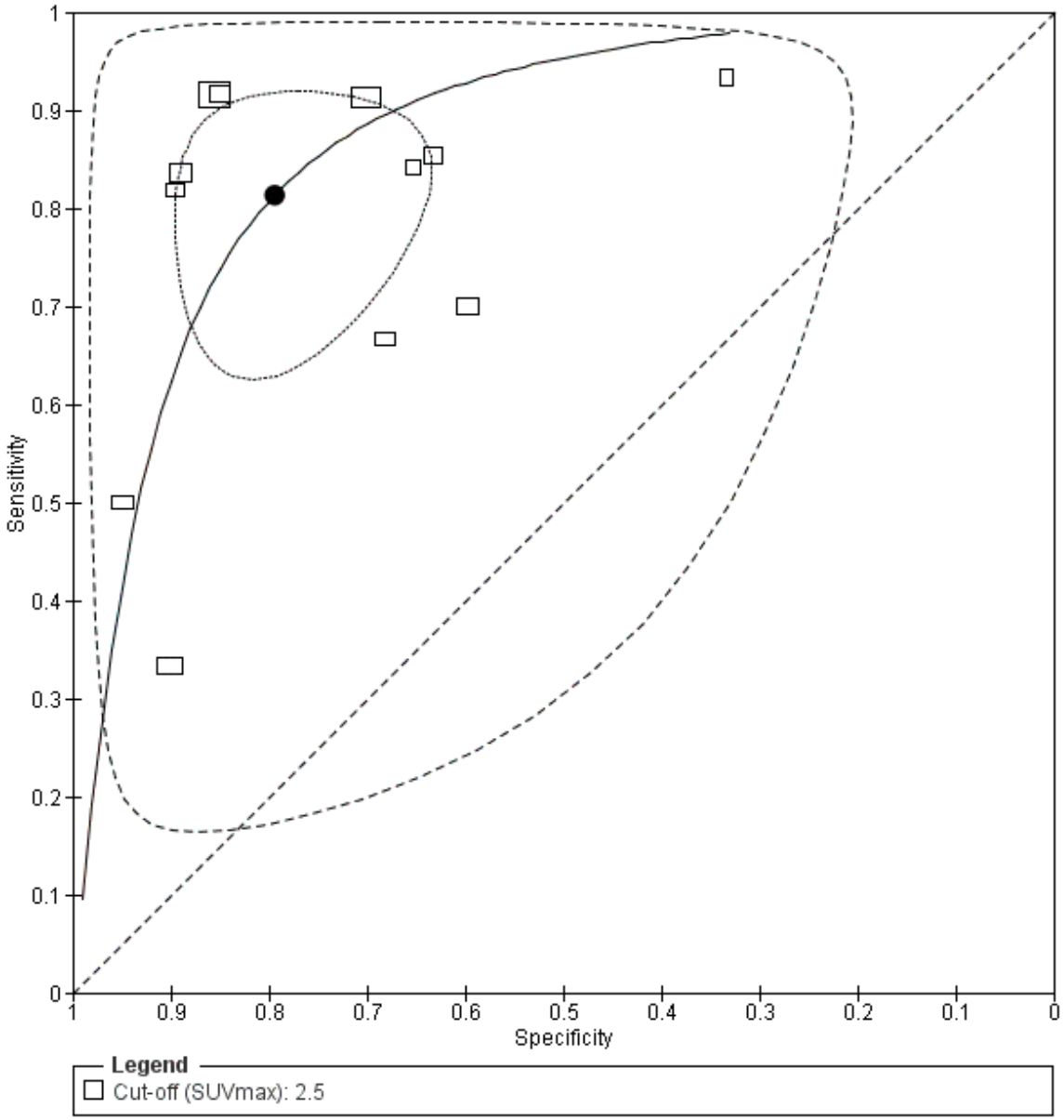
We conducted two primary analyses based on the criteria for test positivity: Regarding the *Activity* > *background* group, 18 of the included studies employed a qualitative criterion for test positivity based on the relative activation between the lymph nodes and the surrounding tissue (Bille 2013; Chen 2010; De Wever 2007; Gunluoglu 2011; Harders 2012; Jeon 2010; Kim 2007; Kuo 2012; Lee 2007; Lee 2009a; Li 2012a; Plathow 2008; Shin 2008; Tasci 2010; Toba 2010; Yang 2008; Yang 2010; Yi 2008). The summary sensitivity and specificity estimates for this criterion for test positivity were 77.4% (95% CI 65.3 to 86.1) and 90.1% (95% CI 85.3 to 93.5), respectively. Figure 7 shows the accuracy estimates of these studies in ROC space along with a summary ROC curve fitted with the HSROC model. The wide area of the prediction region illustrates that between-study heterogeneity is still high, and the 95% confidence region around the summary value of sensitivity and specificity is also relatively large, denoting lack of precision.

Figure 7. Summary ROC Plot of studies with Activity > background as the criterion for test positivity. Empty squares represent individual study estimates, with the size of the square proportional to the study sample size. The solid line represent the SROC curve. The filled circle is the summary point representing the average sensitivity and specificity estimates. The ellipses around this summary point are the 95% confidence region (dotted line) and the 95% prediction region (dashed line). The dashed upward diagonal represents the completely uninformative test



Regarding the *SUVmax ≥ 2.5* group, 12 studies used a common cut-off value of SUVmax of ≥ 2.5 (Bryant 2006a; Hu 2011; Hwangbo 2009; Iskender 2012; Koksal 2013; Li 2010; Ose 2012; Sanli 2009; Saydam 2012; Subedi 2009; Uruga 2011; Uskul 2009). The summary sensitivity and specificity estimates for this most common threshold were 81.3% (95% CI 70.2 to 88.9) and 79.4% (95% CI 70 to 86.5), respectively. Figure 8 shows the accuracy estimates of these studies in ROC space, along with a summary ROC curve fitted with the HSROC model. The wide area of the prediction region illustrates that between-study heterogeneity is very high. The 95% confidence region around the summary value of sensitivity and specificity is also large, denoting a clear lack of precision.

Figure 8. Summary ROC Plot of studies with SUVmax > 2.5 as the criterion for test positivity. Empty squares represent individual study estimates, with the size of the square proportional to the study sample size. The solid line represent the SROC curve. The filled circle is the summary point representing the average sensitivity and specificity estimates. The ellipses around this summary point are the 95% confidence region (dotted line) and the 95% prediction region (dashed line). The dashed upward diagonal represents the completely uninformative test

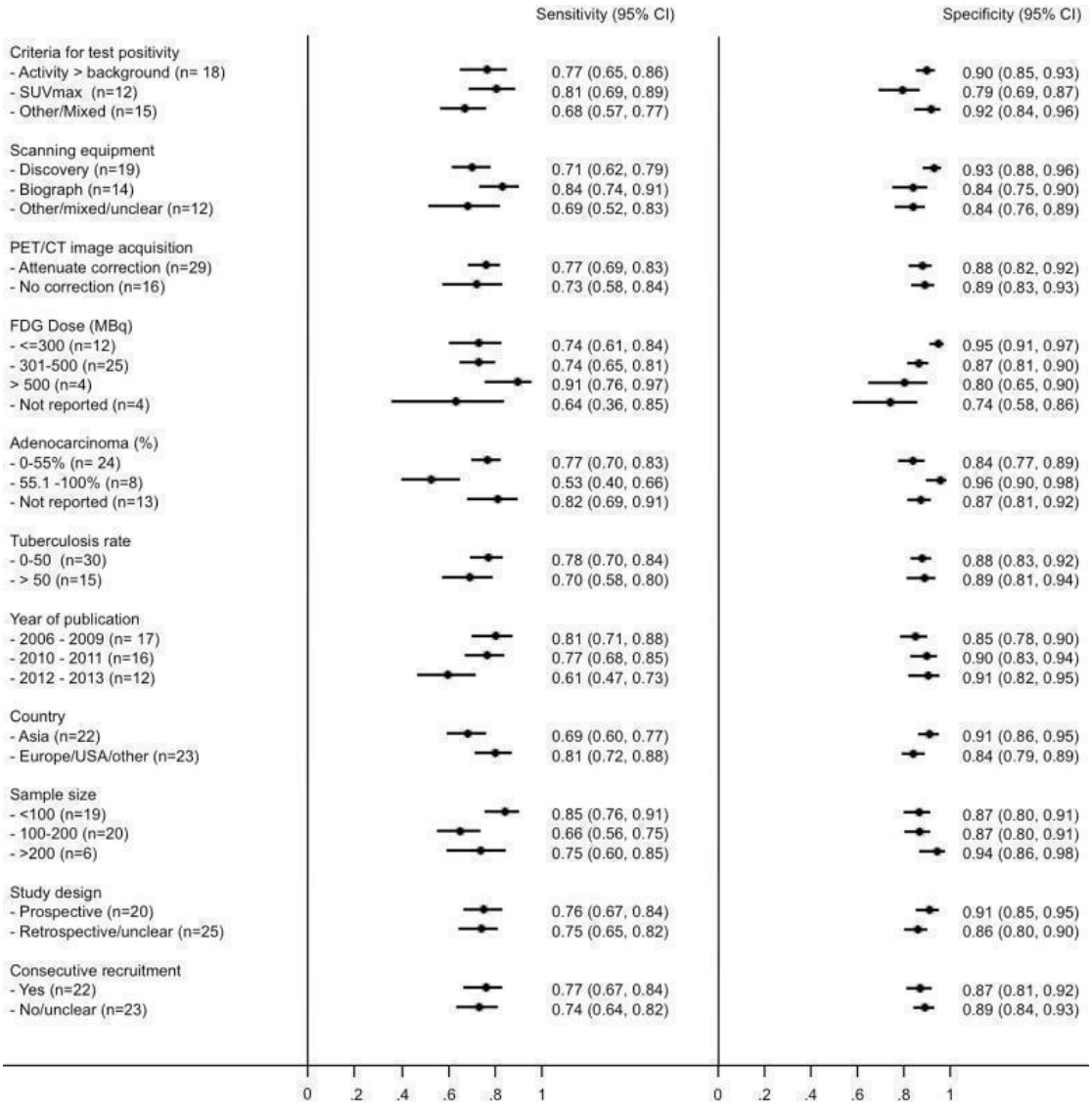


We did not conduct any further analyses for the studies using *Other/mixed/unclear* criteria for test positivity as the large variation in these criteria would have made any further analyses meaningless.

Investigations of heterogeneity

We report on subgroups based on preplanned covariates that were anticipated to contribute to heterogeneity. We did not conduct these analyses separately for the two main analyses as we found no overall effect of test positivity criteria (*Activity > background* (sensitivity = 0.77, 95% CI 0.65 to 0.86; specificity = 0.9, 95% CI 0.85 to 0.93) versus *SUVmax ≥ 2.5* (sensitivity = 0.81, 95% CI 0.69 to 0.89; specificity = 0.79, 95% CI 0.69 to 0.87) versus *Other/mixed* (sensitivity = 0.68, 95% CI 0.57 to 0.77; specificity = 0.92, 95% CI 0.84 to 0.96); P = 0.77). Figure 9 presents the results of the subgroup analysis.

Figure 9. Investigations of possible sources of heterogeneity



In summary, there were differences in accuracy for a number of factors (country, type of PET-CT scanner, percentage of participants with adenocarcinoma, FDG dose, and study size), while we observed no differences for design characteristics, year of publication, and tuberculosis incidence rate. The detailed results for each factor follow.

Country of origin was significantly associated with diagnostic accuracy. Studies performed in western countries (Europe/USA/other: sensitivity = 0.81, 95% CI 0.72 to 0.88; specificity = 0.84, 95% CI 0.79 to 0.89) showed greater sensitivity (p (pair-wise) = 0.045) and lower specificity (P (pair-wise) = 0.035) than studies performed in Asian countries (sensitivity = 0.69, 95% CI 0.6 to 0.77; specificity = 0.91, 95% CI 0.86 to 0.95)); P (overall effect) = 0.04). The type of PET-CT scanner was also associated with different diagnostic accuracy: Compared to Discovery (sensitivity = 0.71, 95% CI 0.62 to 0.79; specificity = 0.93, 95% CI 0.88 to 0.96), Biograph scanning equipment (sensitivity = 0.84, 95% CI 0.74 to 0.91; specificity = 0.84, 95% CI 0.75 to 0.9) showed greater sensitivity (P (pair-wise) = 0.039) and lower specificity (P (pair-wise) = 0.047; P (overall effect) = 0.039). Thirty-two studies reported the percentage of participants with adenocarcinoma and

ranged from 20.5% to 87.2%. There was a clear split in the data between 54.8% and 69.1%, and we therefore employed a cut-off of 55% to analyse this covariate with three levels: 0% to 55% (in effect, this is 20.5% to 54.8%; N = 24), 55.1% to 100% (in effect, 69.1% to 87.2%; N = 8), and not reported (N = 13).

The percentage of participants with adenocarcinoma also influenced the diagnostic accuracy of PET-CT (P (overall effect) = 0.003) with the sensitivity being significantly higher (P (pair-wise) = 0.004) and specificity (P (pair-wise) = 0.001) being significantly lower in studies with ≤ 55% adenocarcinoma participants (sensitivity = 0.77, 95% CI 0.7 to 0.83; specificity = 0.84, 95% CI 0.77 to 0.89) compared with studies with > 55% adenocarcinoma participants (sensitivity = 0.53, 95% CI 0.4 to 0.66; specificity = 0.96, 95% CI 0.9 to 0.98). We also found that FDG dose was associated with different diagnostic accuracy estimates (P (overall effect) = 0.015): Sensitivity was significantly higher (P (pair-wise) = 0.031) and specificity significantly lower (P (pair-wise) = 0.044) in studies using > 500 MBq (sensitivity = 0.91, 95% CI 0.76 to 0.97; specificity = 0.8, 95% CI 0.65 to 0.9) compared with studies using 300 or less MBq (sensitivity = 0.74, 95% CI 0.61 to 0.84; specificity = 0.95, 95% CI 0.91 to 0.97); specificity was also significantly lower (P (pair-wise) = 0.003) in studies using 301 to 500 MBq (sensitivity = 0.74, 95% CI 0.65 to 0.81; specificity = 0.87, 95% CI 0.81 to 0.9) compared with studies using 300 or less MBq, and sensitivity was significantly higher in studies using > 500 MBq compared with those using 301 to 500 MBq (P (pair-wise) = 0.007). The heterogeneity analyses revealed one final covariate that influenced sensitivity and specificity, namely, study size (P (overall effect) = 0.025) with significantly higher sensitivity in studies with < 100 participants (sensitivity = 0.85, 95% CI 0.76 to 0.91; specificity = 0.87, 95% CI 0.8 to 0.91) compared with studies with 100 to 199 participants ((sensitivity = 0.66, 95% CI 0.56 to 0.75; specificity = 0.87, 95% CI 0.8 to 0.91); P (pair-wise) = 0.003) and significantly higher specificity in studies with > 200 participants (sensitivity = 0.75, 95% CI 0.6 to 0.85; specificity = 0.94, 95% CI 0.86 to 0.98) compared with studies with < 100 participants (P (pair-wise) = 0.045) and studies with 100 to 199 participants (P (pair-wise) = 0.0495).

No other analysed covariates were associated with different diagnostic accuracy of the test: Design (prospective (sensitivity = 0.76, 95% CI 0.67 to 0.84; specificity = 0.91, 95% CI 0.85 to 0.95) versus retrospective/unclear (sensitivity = 0.75, 95% CI 0.65 to 0.82; specificity = 0.86, 95% CI 0.8 to 0.9); P = 0.444), consecutive recruitment (yes (sensitivity = 0.77, 95% CI 0.67 to 0.84; specificity = 0.87, 95% CI 0.81 to 0.92) versus no/unclear (sensitivity = 0.74, 95% CI 0.64 to 0.82; specificity = 0.89, 95% CI 0.84 to 0.93); P = 0.933), attenuation correction (yes (sensitivity = 0.77, 95% CI 0.69 to 0.83; specificity = 0.88, 95% CI 0.82 to 0.92) versus no/unclear (sensitivity = 0.73, 95% CI 0.58 to 0.84; specificity = 0.89, 95% CI 0.83 to 0.93); P = 0.55), year of publication (2006 to 2009 (sensitivity = 0.81, 95% CI 0.71 to 0.88; specificity = 0.85, 95% CI 0.78 to 0.9) versus 2010 to 2011 (sensitivity = 0.77, 95% CI 0.68 to 0.85; specificity = 0.9, 95% CI 0.83 to 0.94) versus 2012 to 2013 (sensitivity = 0.61, 95% CI 0.47 to 0.73; specificity = 0.91, 95% CI 0.82 to 0.95); P = 0.139), and tuberculosis incidence rate per 100,000 population (0 to 50 (sensitivity = 0.78, 95% CI 0.7 to 0.84; specificity = 0.88, 95% CI 0.83 to 0.92) versus > 50 (sensitivity = 0.7, 95% CI 0.58 to 0.8; specificity = 0.89, 95% CI 0.81 to 0.94); P = 0.688).

Where the overall effect of the covariate was significant but one of the levels of the covariate was not reported (adenocarcinoma), *Other/mixed/unclear* (scanning equipment), or unclear (FDG dose), we have not reported any pair-wise comparisons involving that level of the covariate because we would not be able to make any useful statements about such analyses.

Sensitivity analysis

In Table 3, we present this restricted analysis including only studies with low risk of bias or low concerns about applicability. Table 3 seems to suggest that in the *Activity* > *background* group, the overall estimate of sensitivity especially is sensitive to selection bias; reference standard bias; and clear definition of test positivity; and to a lesser extent, index test bias and commercial funding bias, with lower combined estimates of sensitivity observed for all the low 'Risk of bias' studies compared with the full analysis. In the *SUVmax* ≥ *2.5* group, the sensitivity analyses suggest that both overall accuracy estimates are much less sensitive to the exclusion of studies according to the covariates analysed. Only flow and timing bias and commercial funding bias led to slightly lower estimates of both sensitivity and specificity. We did not make any formal statistical comparison given the scarce number of studies analysed after the exclusions in the sensitivity analysis.

Summary of findings

PET-CT for assessing mediastinal lymph node involvement in participantswith suspected resectable non-small cell lung cancer					
Population	Participants with suspected/confirmed NSCLC who are considered potentially suitable for primary resection				
Index test	PET-CT carried out on the various available integrated PET-CT scanners with cut-off values for test positivity as reported in the included studies				
Target condition	Resectable NSCLC defined as NSCLC that has not spread to either the ipsilateral mediastinal lymph nodes, subcarinal (N2) lymph nodes, or both				
Reference standard	Pathological confirmation of PET-CT results				
Included studies	45 studies with 6095 participants available for analysis (median = 112, interquartile range (IQR) = 54 to 169), 4551 of whom were N0 or N1 and 1544 participants were N2 or N3 Different criteria for test positivity were used in the included studies: *Activity > background* (18 studies; N = 2823; prevalence of N2 and N3 nodes = 679/2328) *SUVmax ≥ 2.5* (12 studies; N = 1656; prevalence of N2 and N3 nodes = 465/1656) *Other/mixed* criteria for test positivity (15 studies; N = 1616; prevalence of N2 and N3 nodes = 400/1616) None of the studies reported (any) adverse events				
Test subgroup	**Number of participants (studies)**	**Prevalence %**	**Summary accuracy % (95% CI)**	**Implications**	**Quality and comments**
Activity > background	2823 (18)	29.2	Sensitivity: 77.4 (65.3 to 86.1) Specificity: 90.1 (85.3 to 93.5)	With the observed prevalence, there will be 66 missed cases and 70 cases who will receive futile surgery	Participant selection, index test, and flow and timing poorly reported Population spectrum narrower than in standard clinical practice in a substantial number of studies Results sensitive to selection bias, reference standard bias, and clear definition of test positivity Substantial heterogeneity was observed

SUVmax $\geq$ 2.5	1656 (12)	28.1	Sensitivity: 81.3 (70.2 to 88.9) Specificity: 79.4 (70.0 to 86.5)	With the observed prevalence, there will be 53 missed cases and 148 cases who will receive futile surgery	Participant selection, index test and flow, and timing poorly reported Population spectrum narrower than in standard clinical practice in a substantial number of studies Results sensitive to flow and timing bias and commercial funding bias Substantial heterogeneity was observed
All included studies	6095	25.3	Heterogeneity analyses showed significant contributions to between-study heterogeneity from the following covariates: country of study origin, percentage of participants with adenocarcinoma, FDG dose, type of PET-CT scanner, and study size. Study design, consecutive recruitment, attenuation correction, year of publication, and tuberculosis incidence rate per 100,000 population did not contribute significantly to the observed heterogeneity		
CAUTION: The results in this table should **not** be interpreted in isolation from the results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the review					

CI = confidence interval.

FDG = (^{18} F)-2-fluoro-deoxy-D-glucose.
IQR = interquartile range.
NSCLC = non-small cell lung cancer.
PET-CT = positron emission tomography-computed tomography.
SUVmax = maximum standardised uptake value.

DISCUSSION

We have conducted an up-to-date review of studies that have examined the role of PET-CT in determining whether there has been N2 or N3 disease. This is important because it is often crucial in planning treatment. People with N2 or N3 disease do not usually undergo radical treatment with surgery as their primary treatment and instead receive palliative treatment. Where radical treatment is to be offered, there needs to be careful planning, and knowledge of the status of N2 or N3 nodes is essential.

Summary of main results

We found that there was considerable variation in sensitivity and specificity amongst the 45 studies evaluated. Our two main analyses showed that in the studies employing a criterion of *Activity > background*, the summary sensitivity and specificity estimates were 77.4% (95% CI 65.3 to 86.1) and 90.1% (95% CI 85.3 to 93.5), respectively, and for studies employing *SUVmax ≥ 2.5* as the criterion for test positivity, the sensitivity and specificity estimates for this threshold were 81.3% (95% CI 70.2 to 88.9) and 79.4% (95% CI 70 to 86.5), respectively. However, it was the case for both analyses that the prediction and confidence regions were large, and further analyses found that the following covariates partly explained between-study variability: country of origin, with studies performed in western countries showing greater sensitivity and lower specificity than studies performed in Asian countries; type of PET-CT scanner, with Biograph scanning equipment showing greater sensitivity and lower specificity than Discovery; the percentage of participants with adenocarcinoma, with the sensitivity being significantly higher and specificity significantly lower in studies with ≤ 55% adenocarcinoma participants compared with studies with > 55% adenocarcinoma participants; FDG dose, with significantly higher sensitivity and significantly lower specificity in studies using > 500 MBq compared with studies using 300 or less MBq, significantly lower specificity in studies using 301 to 500 MBq compared with studies using 300 or less MBq, and significantly higher sensitivity in studies using > 500 MBq compared with those using 301 to 500 MBq; and study size, with significantly higher sensitivity in studies with < 100 participants compared with studies with 100 to 199 participants and significantly higher specificity in studies with 200+ participants compared with studies with > 100 participants and studies with 100 to 199 participants. Sensitivity analyses also suggested that the summary estimates from the two main analyses were sensitive to a number of biases. Specifically, in the *Activity > background* group, the overall estimate of sensitivity especially is sensitive to selection bias; reference standard bias; clear definition of test positivity; and to a lesser extent, index test bias and commercial funding bias, with lower combined estimates of sensitivity observed for all the low 'Risk of bias' studies compared with the full analysis. In the *SUVmax ≥ 2.5* group, the sensitivity analyses suggested that both overall accuracy estimates were somewhat sensitive to flow and timing bias and commercial funding bias, which led to slightly lower estimates of both sensitivity and specificity.

The observation that studies performed in western countries showed higher sensitivity and lower specificity compared with studies performed in Asian countries may be linked to the observation that the sensitivity was significantly higher and specificity significantly lower in studies with ≤ 55% adenocarcinoma participants compared with studies with > 55% adenocarcinoma participants. This is because we know that there are differences in tumour biology of lung cancer in east Asians, with a greater proportion of cancers being adenocarcinoma and in non-smokers (Maemondo 2010). This may influence the FDG uptake, which is lower in adenocarcinoma than in other common forms of NSCLC (Casali 2010; Davidson 2009; Jeong 2002; Lu 2010), and therefore the sensitivity, which also fits well with the finding that studies using a relatively higher dose of FDG (> 500 MBq) had a higher sensitivity than those using a relatively lower dose (< 500 MBq).

The observation that the type of PET-CT scanner employed is associated with different accuracy estimates suggests that the two main integrated PET-CT scanner manufacturers have produced products with different characteristics, and it is of potential importance to lung cancer clinicians to know that the equipment alone may influence the result obtained.

The finding that study size and various biases influenced the results of this review underscores the need for well-designed and adequately powered (and reported) diagnostic test accuracy studies in NSCLC staging research, specifically, but also in diagnostic medical research in general. It should however also be noted that even within the different test positivity criteria subgroups, the actual criteria/cut-offs used varied between the studies (e.g., in the *Other/mixed/unclear* group, SUVmax > 3.5 versus ≥ 4.1 versus ≥ 4.45), which may well explain a significant amount of the remaining between-study heterogeneity. Unfortunately, we were unable to investigate the contribution of this variable in greater detail because of the low number of studies within each test positivity subcategory. We also note that SUVmax may show some variation in value on repeated measurements, but this is minimal in relation to the spread of values normally obtained in studies (i.e., few are on the cut-off value). There may be some variation in the SUVmax measured in different centres, but this is again likely to be minimal as the majority of the measurement is standardised by the software (Lindholm 2014).

Strengths and weaknesses of the review

We performed an extensive search for relevant studies and were able to obtain data from 45 studies for inclusion. With these studies, we were able to show that a number of conceivably connected factors influence the accuracy of PET-CT for mediastinal staging of NSCLC, namely adenocarcinoma; Asian population; and FDG

dose, as well as by scanner type, a finding which we do not believe is linked to the influence of the other covariates after careful examination of potential overlap between the studies that contributed to the different results. However, despite the relatively large number of relevant studies and a number of prespecified heterogeneity and sensitivity analyses, a substantial amount of unexplained heterogeneity still mark the results, which we hypothesise can, at least in part, be explained by the large variation in the criteria used for test positivity in the different studies. Unfortunately, we were unable to examine in detail this hypothesis because too few studies used the same criteria.

While we are also reasonably confident that the reference standard used in our review is robust and clinically appropriate, it should be noted that some of these tests themselves have limitations in their accuracy. Where EBUS-TBNA was positive, there was often no further sampling. False positives would be very uncommon and unlikely to influence the results unless there was a systematic error within the study by the clinicians involved. Where EBUS-TBNA or mediastinoscopy was negative, reliance was placed either on a period of follow-up, confirming negativity, or systematic nodal dissection and sampling as part of surgery. Where the latter was not clearly specified or the nodal sampling potentially was incomplete, we identified this as a potential source of bias. However, we were clearly unable to assess the quality of adherence to the protocol specified in the studies. It is possible that nodal sampling quality varied amongst surgeons and studies, although most of these studies were conducted at large centres where one would expect high standards. There is also a small risk that N3 nodes might have been missed where mediastinoscopy was not performed prior to surgery. This would mean that contralateral nodes would only have been sampled by EBUS-TBNA, as only ipsilateral nodes are sampled in a systematic nodal dissection. A further limitation is that we were unable to find sufficient studies that looked at the accuracy of PET-CT in lymph nodes that were not significantly enlarged by CT criteria ($\leq$ 10 mm maximum short axis diameter). However, those studies with a low prevalence of malignancy were likely to have included people with smaller nodes as nodal size is strongly correlated with malignancy; in these studies, specificity appeared to be high. Lastly, we would have preferred to be able to include more studies in potentially more difficult populations, such as those with a high prevalence of diseases or conditions known to produce false positive results, such as tuberculosis and industrial exposure to pathogens. Unfortunately, this was not possible as not enough relevant studies appear to have been conducted in such populations.

Applicability of findings to the review question

Broadly speaking, our findings are applicable to the review question in terms of the index test and reference standard where, generally, there was good correspondence between the tests used in the included studies and those specified in our review question. However, as outlined in Methodological quality of included studies, a substantial number of the included studies only included participants who had received resection for NSCLC while other studies only included participants with T1 NSCLC or who were retired coal workers. And all of these inclusion restrictions artificially narrow the range of patients who would receive FDG/PET-CT in standard practice, in particular, the patients with N2 or N3 disease, and this, in turn, gives rise to high concerns about the applicability of the populations to the question of the present review. On the other hand, enough studies were available to enable us to address these applicability concerns through sensitivity analyses, which suggested that sensitivity is increased while specificity is decreased relative to the overall estimates within both of the main analyses when we only analysed the studies with low concerns about applicability. We believe that these results are directly applicable to the typical populations seen in routine clinical practice (accepting that these differ in different countries). We have shown clearly that there is variation in the accuracy of PET-CT in the differentiation of N2 and N3 lymph node metastasis in NSCLC and that this variation is related to, among other factors, NSCLC subtype (adenocarcinoma), country of study origin (Asia), FDG dose, and PET-CT scanner type, all of which should be born in mind by the lung cancer diagnostician.

AUTHORS' CONCLUSIONS

Implications for practice

This review has provided up-to-date data on the accuracy of PET-CT scanning in determining N2 and N3 nodal status in non-small cell lung cancer. It has shown that pooled sensitivity and specificity, whilst reasonable at around 0.8, is insufficient to allow management based on PET-CT alone. In clinical practice, PET-CT is a useful test, and this review supports that. However, the review has also confirmed that PET-CT has to form part of a clinical pathway supported by other investigations and cannot be used as a stand-alone test. The findings therefore support NICE guidance on this topic, where PET-CT is used to guide clinicians in the next step, which is either a biopsy or where negative and nodes are small, directly to surgery (NICE 2011). The apparent difference between the two main makes of PET-CT scanner is important, as this appears independent of the operator or other factors. This a new finding, to our knowledge, and may be important for lung cancer multidisciplinary teams to know. The relatively low sensitivity but high specificity of the Discovery could, in some circumstances, such as where the patient is of very borderline fitness, influence the decision. This would, as is recommended, include the wishes of the patient after a fully informed discussion. The difference between makes and the general variability of results suggests that all large centres should actively monitor their accuracy so that they can make reliable decisions based on their own results.

The pooled results by country identified important differences in the accuracy of PET-CT, showing that it may be less sensitive in Asian countries. Again, this calls for centres to audit their results and identify the populations in which PET-CT is of most use or potentially little value.

Implications for research

In radiology, as in many other areas of medicine, technological advances may lead to rapid changes in clinical practice. Newer PET-CT scanners will be introduced that have higher resolution and lower radiation dose. As they are expensive, it will take some time before they become universally used, but it will be important to measure their accuracy as soon as possible. A key question will be how they perform in different populations and according to the size of lymph nodes. Studies should be designed in populations with a high prevalence of tuberculosis or industrial disease and in participants with interstitial lung disease. These patients are commonly encountered in clinical practice, and we do not know exactly how these conditions alter the accuracy of PET-CT. There should be correspondence between the protocols and make of scanners in studies conducted in Asia and in the western countries so that comparisons can be made and so that populations can be identified where PET-CT is of use or no use. The reasons for our observed difference between the make of PET-CT scanners are not clear, and studies should be undertaken to establish the reason for the difference, which is likely to relate to calibration rather than a difference in the accuracy of detectors. Another key question is whether some N2 nodes that are shown to be positive on PET-CT, with or without pathological confirmation, should be resected as part of a definitive operation. This may depend on nodal size, SUV, and number of nodal stations involved. NICE clinical guideline 121 (NICE 2011) has already made recommendations for research into this area.

ACKNOWLEDGEMENTS

We would like to thank the National Institute for Health and Care Excellence (NICE), the National Collaborating Centre for Cancer, and the Guideline Development Group for the NICE clinical guideline 121 (The diagnosis and treatment of lung cancer (update)) (NICE 2011). We would also like to thank all the peer reviewers who have provided us with helpful comments on earlier versions of this review (and the protocol before that) and the Cochrane Lung Cancer group for all their continued advice and assistance.

Marta Roqué i Fíguls is a PhD candidate at the Department of Paediatrics, Obstetrics and Gynecology and Preventive Medicine, Universitat Autònoma de Barcelona, Spain.

REFERENCES

References to studies included in this review

Bille 2013 ***{published data only}***

Billè A, Okiror L, Skanjeti A, Errico L, Arena V, Penna D, et al.Evaluation of integrated positron emission tomography and computed tomography accuracy in detecting lymph node metastasis in patients with adenocarcinoma vs squamous cell carcinoma. *European Journal of Cardio-Thoracic Surgery* 2013;**43**(3):574–9. [PUBMED: 22689182]

Billè A, Okiror L, Skanjeti A, Errico L, Arena V, Penna D, et al.The prognostic significance of maximum standardized uptake value of primary tumor in surgically treated non-small-cell lung cancer patients: analysis of 413 cases. *Clinical Lung Cancer* 2013;**14**(2):149–56. [PUBMED: 22682667]

Billé A, Pelosi E, Skanjeti A, Arena V, Errico L, Borasio P, et al.Preoperative intrathoracic lymph node staging in patients with non-small-cell lung cancer: accuracy of integrated positron emission tomography and computed tomography. *European Journal of Cardio-Thoracic Surgery* 2009;**36**(3): 440–5. [PUBMED: 19464906]

Bryant 2006a ***{published data only}***

Bryant AS, Cerfolio RJ, Klemm KM, Ojha B. Maximum standard uptake value of mediastinal lymph nodes on integrated FDG-PET-CT predicts pathology in patients with non-small cell lung cancer. *The Annals of Thoracic Surgery* 2006;**82**(2):417–22. [PUBMED: 16863739]

Cerfolio RJ, Bryant AS, Ojha B, Eloubeidi M. Improving the inaccuracies of clinical staging of patients with NSCLC: a prospective trial. *The Annals of Thoracic Surgery* 2005;**80** (4):1207–13. [PUBMED: 16181842]

Carnochan 2009 ***{published data only}***

Carnochan FM, Walker WS. Positron emission tomography may underestimate the extent of thoracic disease in lung cancer patients. *European Journal of Cardio-Thoracic Surgery* 2009;**35**(5):781–4. [PUBMED: 19272791]

Chen 2010 ***{published data only}***

Chen W, Jian W, Li HT, Li C, Zhang YK, Xie B, et al.Whole-body diffusion-weighted imaging vs. FDG-PET for the detection of non-small-cell lung cancer. How do they measure up?. *Magnetic Resonance Imaging* 2010;**28**(5): 613–20. [PUBMED: 20418042]

Czepczynski 2011 ***{published data only}***

Czepczynski R, Zielinski P, Stangierski A, Gabryel P, Kasprzak W, Dyszkiewicz W, et al.The value of 18F-FDG PET/CT scan in the evaluation lymph node status in patients with non-small cell lung carcinoma. *European*

Journal of Nuclear Medicine and Molecular Imaging 2011; **Conference**:S282.

Darling 2011 *{published data only}*

Darling GE, Maziak DE, Inculet RI, Gulenchyn KY, Driedger AA, Ung YC, et al.Positron emission tomography-computed tomography compared with invasive mediastinal staging in non-small cell lung cancer: results of mediastinal staging in the early lung positron emission tomography trial. *Journal of Thoracic Oncology* 2011;**6**(8):1367–72. [PUBMED: 21587082]

Maziak DE, Darling GE, Inculet RI, Gulenchyn KY, Driedger AA, Ung YC, et al.Positron emission tomography in staging early lung cancer: a randomized trial. *Annals of Internal Medicine* 2009;**151**(4):221–8. [PUBMED: 19581636]

De Wever 2007 *{published data only}*

De Wever W, Ceyssens S, Mortelmans L, Stroobants S, Marchal G, Bogaert J, et al.Additional value of PET-CT in the staging of lung cancer: comparison with CT alone, PET alone and visual correlation of PET and CT. *European Radiology* 2007;**17**(1):23–32. [PUBMED: 16683115]

El-Hariri 2012 *{published data only}*

El-Hariri MA, Gouhar GK, Refat AM. Integrated PET/CT in the preoperative staging of lung cancer: a prospective comparison of CT, PETand integrated PET/CT. *The Egyptian Journal of Radiology and Nuclear Medicine* 2012;**43**(4):613–21. [DOI: 10.1016/j.ejrnm.2012.09.007]

Fischer 2011 *{published data only}*

Fischer B, Lassen U, Mortensen J, Larsen S, Loft A, Bertelsen A, et al.Preoperative staging of lung cancer with combined PET-CT. *The New England Journal of Medicine* 2009;**361**(1):32–9. [PUBMED: 19571281]

Fischer BM, Loft A, Bertelsen AK, Mortensen J, Lassen U. Diagnostic accuracy of PET/CT assigning overall TNM-stage in patients with NSCLC according to the 1997 respectively the 2010 TNM classification system. *European Journal of Nuclear Medicine and Molecular Imaging* 2011; **Conference**:S123.

Fischer BM, Mortensen J, Hansen H, Vilmann P, Larsen SS, Loft A, et al.Multimodality approach to mediastinal staging in non-small cell lung cancer. Faults and benefits of PET-CT: a randomised trial. *Thorax.* 2011;**66**(4):294–300. [PUBMED: 21169287]

Gunluoglu 2011 *{published data only}*

Gunluoglu MZ, Melek H, Medetoglu B, Demir A, Kara HV, Dincer SI. The validity of preoperative lymph node staging guidelines of European Society of Thoracic Surgeons in non-small-cell lung cancer patients. *European Journal of Cardio-thoracic Surgery* 2011;**40**(2):287–90. [PUBMED: 21185733]

Harders 2012 *{published data only}*

Harders SW. LUCIS: lung cancer imaging studies. *Danish Medical Journal* 2012;**59**(11):B4542. [PUBMED: 23171752]

Harders, SW, Madsen HH, Hjorthaug K, Arveschoug AK, Rasmussen TR, Meldgaard, P, et al.Mediastinal staging in non-small cell lung carcinoma: CT versus F-18-FDG PET/CT. Cancer Imaging in press.

Hu 2011 *{published data only}*

Hu M, Han A, Xing L, Yang W, Fu Z, Huang C, et al.Value of dual-time-point FDG PET/CT for mediastinal nodal staging in non-small-cell lung cancer patients with lung comorbidity. *Clinical Nuclear Medicine* 2011;**36**(6):429–33. [PUBMED: 21552018]

Hu M, Yu J, Xing L, Han A, Kong L. Diagnostic ability of dual-time-point fdg PET/CT for mediastinal lymph node metastases in non-small cell lung cancer patients. *International Journal of Radiation Oncology Biology Physics* 2010;**78**(3 Suppl):S500. [DOI: 10.1016/j.ijrobp.2010.07.1169]

Hwangbo 2009 *{published data only}*

Hwangbo B, Kim SK, Lee HS, Lee HS, Kim MS, Lee JM, et al.Application of endobronchial ultrasound-guided transbronchial needle aspiration following integrated PET/CT in mediastinal staging of potentially operable non-small cell lung cancer. *Chest* 2009;**135**(5):1280–7. [PUBMED: 19118267]

Iskender 2012 *{published data only}*

Iskender I, Kadioglu SZ, Cosgun T, Kapicibasi HO, Sagiroglu G, Kosar A, et al.False-positivity of mediastinal lymph nodes has negative effect on survival in potentially resectable non-small cell lung cancer. *European Journal of Cardio-Thoracic Surgery* 2012;**41**(4):874–9. [PUBMED: 22423060]

Iskender I, Kadioglu SZ, Kosar A, Atasalihi A, Kir A. Is there any maximum standardized uptake value variation among positron emission tomography scanners for mediastinal staging in non-small cell lung cancer?. *Interactive Cardiovascular and Thoracic Surgery* 2011;**12**(6):965-9. [PUBMED: 21441257]

Iskender I, Kapicibasi HO, Kadioglu SZ, Sevilgen G, Tezel C, Kosar A, et al.Comparison of integrated positron emission tomography/computed tomography and mediastinoscopy in mediastinal staging of non-small cell lung cancer: analysis of 212 patients. *Acta Chirurgica Belgica* 2012;**112**(3):219–25. [PUBMED: 22808763]

Jeon 2010 *{published data only}*

Jeon TY, Lee KS, Yi CA, Chung MP, Kwon OJ, Kim B-T, et al.Incremental value of PET/CT over CT for mediastinal nodal staging of non-small cell lung cancer: comparison between patients with and without idiopathic pulmonary fibrosis. *American Journal of Roentgenology* 2010;**195**(2):370–6. [PUBMED: 20651192]

Kim 2007 *{published data only}*

Kim BT, Lee KS, Shim SS, Choi JY, Kwon OJ, Kim H, et al.Stage T1 non-small cell lung cancer: preoperative mediastinal nodal staging with integrated FDG PET/CT--a prospective study. *Radiology* 2006;**241**(2):501–9. [PUBMED: 16966480]

Kim YK, Lee KS, Kim BT, Choi JY, Kim H, Kwon OJ, et al.Mediastinal nodal staging of nonsmall cell lung cancer using Integrated 18F-FDG PET/CT in a tuberculosis-

endemic country: diagnostic efficacy in 674 patients. *Cancer* 2007;**109**(6):1068–77. [PUBMED: 7311309]

Koksal 2013 ***{published data only}***
Koksal D, Demirag F, Bayiz H, Ozmen O, Tatci E, Berktas B, et al.The correlation of SUVmax with pathological characteristics of primary tumor and the value of tumor/lymph node SUVmax ratio for predicting metastasis to lymph nodes in resected NSCLC patients. *Journal of Cardiothoracic Surgery* 2013;**8**:63–70. [PUBMED: 23557204]

Kuo 2012 ***{published data only}***
Kuo WH, Wu YC, Wu CY, Ho KC, Chiu PH, Wang CW, et al.Node/aorta and node/liver SUV ratios from (18)F-FDG PET/CT may improve the detection of occult mediastinal lymph node metastases in patients with non-small cell lung carcinoma. *Academic Radiology* 2012;**19**(6):685–92. [PUBMED: 22459646]

Lee 2007 ***{published data only}***
Lee BE, von Haag D, Lown T, Lau D, Calhoun R, Follette D. Advances in positron emission tomography technology have increased the need for surgical staging in non-small cell lung cancer. *The Journal of Thoracis and Cardiovascular Surgery* 2007;**133**(3):746–52. [PUBMED: 17320577]

Lee 2009a ***{published data only}***
Lee JW, Kim BS, Lee DS, Chung JK, Lee MC, Kim S, et al.18F-FDG PET/CT in mediastinal lymph node staging of non-small-cell lung cancer in a tuberculosis-endemic country: consideration of lymph node calcification and distribution pattern to improve specificity. *European Journal of Nuclear Medicine and Molecular Imaging* 2009;**36**(11):1794–802. [PUBMED: 19430783]

Lee 2011 ***{published data only}***
Lee SH, Min JW, Lee CH, Park CM, Goo JM, Chung DH, et al.Impact of parenchymal tuberculosis sequelae on mediastinal lymph node staging in patients with lung cancer. *Journal of Korean Medical Sciences* 2011;**26**(1):67–70. [PUBMED: 21218032]

Lee 2012 ***{published data only}***
Lee SM, Park CM, Paeng JC, Im HJ, Goo JM, Lee HJ, et al.Accuracy and predictive features of FDG-PET/CT and CT for diagnosis of lymph node metastasis of T1 non-small-cell lung cancer manifesting as a subsolid nodule. *European Radiology* 2012;**22**(7):1556–63. [PUBMED: 22358427]

Li 2010 ***{published data only}***
Li XD, Yin JL, Liu WK, Ouyang X, Zhou Z, Qiao GB, et al.[Value of positron emission tomography-computed tomography in the diagnosis of mediastinal lymph node metastasis of non-small cell lung cancer]. *Journal of Southern Medical University* 2010;**30**(3):506–8. [PUBMED: 20335121]

Li 2012a ***{published data only}***
Li M, Wu N, Liu Y, Zheng R, Liang Y, Zhang W, et al.Regional nodal staging with 18F-FDG PET-CT in non-small cell lung cancer: Additional diagnostic value of CT attenuation anddual-time-point imaging. *European Journal of Radiology* 2012;**81**(8):1886–90. [PUBMED: 21511421]

Morikawa 2009 ***{published data only}***
Morikawa M, Demura Y, Ishizaki T, Ameshima S, Miyamori I, Sasaki M, et al.The effectiveness of 18F-FDG PET/CT combined with STIR MRI for diagnosing nodal involvement in the thorax. *Journal of Nuclear Medicine* 2009;**50**(1):81–7. [PUBMED: 19091887]

Ohnishi 2011 ***{published data only}***
Ohnishi R, Yasuda I, Kato T, Tanaka T, Kaneko Y, Suzuki T, et al.Combined endobronchial and endoscopic ultrasound-guided fine needle aspiration for mediastinal nodal staging of lung cancer. *Endoscopy* 2011;**43**(12):1082–89. [PUBMED: 21971924]

Ohno 2011 ***{published data only}***
Ohno Y, Koyama H, Yoshikawa T, Nishio M, Aoyama N, Onishi Y, et al.N stage disease in patients with non-small cell lung cancer: efficacy of quantitative and qualitative assessment with STIR turbo spin-echo imaging, diffusion-weighted MR imaging, and fluorodeoxyglucose PET/CT. *Radiology* 2011;**261**(2):605–15. [PUBMED: 21926377]

Ose 2012 ***{published data only}***
Ose N, Sawabata N, Minami M, Inoue M, Shintani Y, Kadota Y, Okumura M. Lymph node metastasis diagnosis using positron emission tomography with 2-[F-18] fluoro-2-deoxy-D-glucose as a tracer and computed tomography in surgical cases of non-small cell lung cancer. *European Journal of Cardio-Thoracic Surgery* 2012;**42**(1):89–92. [PUBMED: 22290887]

Ozkan 2011 ***{published data only}***
Ozkan EA, Araz M, Soydal C, Aras G. A retrospective analysis of 18F-FDG PET/CT in the primary staging of non-small cell lung cancer (NSCLC). *European Journal of Nuclear Medicine and Molecular Imaging* 2011;**Conference**:S125.

Perigaud 2009 ***{published data only}***
Perigaud C, Bridji B, Roussel JC, Sagan C, Mugniot A, Duveau D, et al.Prospective preoperative mediastinal lymph node staging by integrated positron emission tomography-computerised tomography in patients with non-small-cell lung cancer. *European Journal of Cardio-Thoracic Surgery* 2009;**36**(4):731–6. [PUBMED: 19632852]

Plathow 2008 ***{published data only}***
Plathow C, Aschoff P, Lichy MP, Eschmann S, Hehr T, Brink I, et al.Positron emission tomography/computed tomography and whole-body magnetic resonance imaging in staging of advanced nonsmall cell lung cancer-initial results. *Investigative Radiology* 2008;**43**(5):290–7. [PUBMED: 18424949]

Sanli 2009 ***{published data only}***
Sanli M, Isik AF, Zincirkeser S, Elbek O, Mete A, Tuncozgur B, et al.Reliability of positron emission tomography-computed tomography in identification of mediastinal lymph node status in patients with non-small cell lung cancer. *The Journal of Thoracic and Cardiovascular Surgery* 2009;**138**(5):1200–5. [PUBMED: 19660381]

Saydam 2012 *{published data only}*

Saydam O, Gokce M, Kilicgun A, Tanriverdi O. Accuracy of positron emission tomography in mediastinal node assessment in coal workers with lung cancer. *Medical Oncology* 2012;**29**(2):589–94. [PUBMED: 21380783]

Shin 2008 *{published data only}*

Shin KM, Lee KS, Shim YM, Kim J, Kim BT, Kwon OJ, et al.FDG PET/CT and mediastinal nodal metastasis detection in stage T1 non-small cell lung cancer: prognostic implications. *Korean Journal of Radiology* 2008;**9**(6):481–9. [PUBMED: 19039263]

Sommer 2012 *{published data only}*

Sommer G, Wiese M, Winter L, Lenz C, Klarhofer M, Forrer F, et al.Preoperative staging of non-small-cell lung cancer: comparison of whole-body diffusion-weighted magnetic resonance imaging and 18F-fluorodeoxyglucose-positron emission tomography/computed tomography. *European Radiology* 2012;**22**(12):2859–67. [PUBMED: 22772365]

Subedi 2009 *{published data only}*

Subedi N, Scarsbrook A, Darby M, Korde K, Mc SP, Muers MF. The clinical impact of integrated FDG PET-CT on management decisions in patients with lung cancer. *Lung Cancer* 2009;**64**(3):301–7. [PUBMED: 19004519]

Tasci 2010 *{published data only}*

Tasci E, Tezel C, Orki A, Akin O, Falay O, Kutlu CA. The role of integrated positron emission tomography and computed tomography in the assessment of nodal spread in cases with non-small cell lung cancer. *Interactive Cardiovascular and Thoracic Surgery* 2010;**10**(2):200–3. [PUBMED: 19933240]

Toba 2010 *{published data only}*

Toba H, Kondo K, Otsuka H, Takizawa H, Kenzaki K, Sakiyama S, et al.Diagnosis of the presence of lymph node metastasis and decision of operative indication using fluorodeoxyglucose-positron emission tomography and computed tomography in patients with primary lung cancer. *Journal of Medical Investigation* 2010;**57**(3-4): 305–13. [PUBMED: 20847531]

Tournoy 2007 *{published data only}*

Tournoy KG, Maddens S, Gosselin R, Van Maele G, van Meerbeeck JP, Kelles A. Integrated FDG-PET/CT does not make invasive staging of the intrathoracic lymph nodes in non-small cell lung cancer redundant: a prospective study. *Thorax* 2007;**62**(8):696–701. [PUBMED: 17687098]

Uruga 2011 *{published data only}*

Uruga H. PET/CT for mediastinal lymph node staging in non-small-cell lung cancer with interstitial pneumonia. *American Journal of Respiratory and Critical Care Medicine* 2011;**Conference**:1.

Uskul 2009 *{published data only}*

Üskül BT, Baysungur V, Aksoy F, Turan FE, Sevilgen G, Turker H, et al.Combined use of transbronchial needle aspiration and PET/CT in mediastinal nodal staging of non small cell lung cancer. *Multidisciplinary Respiratory Medicine* 2009;**4**(1):8–14.

Usuda 2013 *{published data only}*

Usuda K, Sagawa M, Motono N, Ueno M, Tanaka M, Machida Y, et al.Advantages of diffusion-weighted imaging over positron emission tomography-computer tomography in assessment of hilar and mediastinal lymph nodes in lung cancer. *Annals of Surgical Oncology* 2013;**20**(5):1676–83. [PUBMED: 23242821]

Usuda K, Zhao XT, Sagawa M, Matoba M, Kuginuki Y, Taniguchi M, et al.Diffusion-weighted imaging is superior to positron emission tomography in the detection and nodal assessment of lung cancers. *The Annals of Thoracic Surgery* 2011;**91**(6):1689–95. [PUBMED: 21619964]

Yang 2008 *{published data only}*

Yang W, Fu Z, Yu J, Yuan S, Zhang B, Li D, et al.Value of PET/CT versus enhanced CT for locoregional lymph nodes in non-small cell lung cancer. *Lung Cancer* 2008;**61**(1): 35–43. [PUBMED: 18177978]

Yang WF, Tan GZ, Fu Z, Yu JM. [Evaluation of the diagnostic value of (18)F-FDG PET-CT and enhanced CT for staging of lymph node metastasis in non small cell lung cancer]. *Zhonghua zhong liu za zhi [Chinese Journal of Oncology]* 2009;**31**(12):925–8. [PUBMED: 20193335]

Yang 2010 *{published data only}*

Yang W, Zhang Y, Fu Z, Yu J, Sun X, Mu D, et al.Imaging of proliferation with 18F-FLT PET/CT versus 18F-FDG PET/CT in non-small-cell lung cancer. *European Journal of Nuclear Medicine and Molecular Imaging* 2010;**37**(7): 1291–9. [PUBMED: 20309686]

Yi 2008 *{published data only}*

Yi CA, Shin, KM, Lee KS, Kim BT, Kim H, et al.Non-small cell lung cancer staging: efficacy comparison of integrated PET/CT versus 3.0-T whole-body MR imaging. *Radiology* 2008;**248**(2):632–42. [PUBMED: 18552311]

References to studies excluded from this review

Agraval 2012 *{published data only}*

Agraval J. Accuracy of FDG-PET/CT nodal staging for non-small-cell lung cancer in surgically resectable cases: a local experience. *Journal of Medical Imaging and Radiation Oncology* 2012;**Conference**:78.

Ahmed 2010 *{published data only}*

Ahmed I, Stuart R, Muller M, Stamenkovic S. PET/CT has enabled the development of a non invasive strategy for mediastinal staging in non small cell lung cancer (NSCLC). *Lung Cancer* 2010;**Conference**:S16. [DOI: 10.1016/S0169-5002(10)70049-0]

Ahmed IM, Stuart R, Stamenkovic S. Impact of NICE guidelines on mediastinal staging strategies. *Lung Cancer* 2010;**Conference**:S16–7. [DOI: 10.1016/S0169-5002(10)70050-7]

Akpinar 2013 *{published data only}*

Akpinar D, Ceylan KC, Duman E, Unsal S, Kaya SO. [The accuracy of positron emission tomography in mediastinal staging of non-small cell lung cancer]. *Türk Göğüs Kalp Damar Cerrahisi Dergisi [Turkish Journal of Thoracic and Cardiovascular Surgery]* 2013;**21**(1):100–5.

Al-Ibraheem 2012 *{published data only}*

Al-Ibraheem A. Performance of FDG PET/CT in mediastinal nodal staging in NSCLC; comparison with chest CE-CT scan and EBUS/TBNA. *European Journal of Nuclear Medicine and Molecular Imaging* 2012;**Conference**: S456.

Allen-Auerbach 2006 *{published data only}*

Allen-Auerbach M, Yeom K, Park J, Phelps M, Czernin J. Standard PET/CT of the chest during shallow breathing is inadequate for comprehensive staging of lung cancer. *Journal of Nuclear Medicine* 2006;**47**(2):298–301. [PUBMED: 16455636]

Al-Sarraf 2008 *{published data only}*

Al-Sarraf N, Gately K, Lucey J, Wilson L, McGovern E, Young V. Lymph node staging by means of positron emission tomography is less accurate in non-small cell lung cancer patients with enlarged lymph nodes: analysis of 1,145 lymph nodes. *Lung Cancer* 2008;**60**(1):62–8. [PUBMED: 17920724]

Al-Sarraf N, Gately K, Lucey J, Wilson L, McGovern E, Young V. Mediastinal lymph node staging by means of positron emission tomography is less sensitive in elderly patients with non-small-cell lung cancer. *Clinical Lung Cancer* 2008;**9**(1):39–43. [PUBMED: 18282357]

An 2008 *{published data only}*

An YS, Sun JS, Park KJ, Hwang SC, Park KJ, Sheen SS, et al.Diagnostic performance of (18)F-FDG PET/CT for lymph node staging in patients with operable non-small-cell lung cancer and inflammatory lung disease. *Lung* 2008;**186**(5):327–36. [PUBMED: 18670805]

Antoch 2003 *{published data only}*

Antoch G, Stattaus J, Nemat AT, Marnitz S, Beyer T, Kuehl H, et al.Non-small cell lung cancer: dual-modality PET/CT in preoperative staging. *Radiology* 2003;**229**(2):526–33. [PUBMED: 14512512]

Aquino 2003 *{published data only}*

Aquino SL, Asmuth JC, Alpert NM, Halpern EF, Fischman AJ. Improved radiologic staging of lung cancer with 2-[18F]-flouro-2-deoxy-D-glucose-positron emission tomography and computed tomography registration. *Journal of Computer Assisted Tomography* 2003;**27**(4):479–84. [PUBMED: 12886128]

BalcıI 2012 *{published data only}*

BalcıI TA. The relationship between Lymph node staging and tumor SUVmax in NSCLC. *European Journal of Nuclear Medicine and Molecular Imaging* 2012;**Conference**: S510.

Beyer 2010 *{published data only}*

Beyer F, Buerke B, Gerss J, Scheffe K, Puesken M, Weckesser M, et al.Prediction of lymph node metastases in NSCLC. Three dimensional anatomical parameters do not substitute FDG-PET-CT. *Nuklearmedizin. Nuclear medicine* 2010;**49**(1):41–8. [PUBMED: 20087533]

Bhatt 2012 *{published data only}*

Bhatt MK. Objective FDG-PET analysis of mediastinal and hilar lymph nodes sampled by endobronchial ultrasound transbronchial needle aspiration. *Internal Medicine Journal* 2012;**Conference**:10.

Booth 2009 *{published data only}*

Booth K, Hanna G, McGonigle N, McGuigan J, McManus K, O'Sullivan J, et al.What is the shelf life of PET/CT staging of the mediastinum in non-small cell lung cancer? . *Interactive Cardiovascular and Thoracic Surgery* 2009; **Conference**:S14–5.

Boulougouri 2012 *{published data only}*

Boulougouri K. Establishing a threshold SUVmax value on 18FDG-PET/CT for identifying metastatic lymph nodes in surgically resectable Non Small Cell Lung Cancer. *European Journal of Nuclear Medicine and Molecular Imaging* 2012; **Conference**:S517.

Bryant 2006b *{published data only}*

Bryant AS, Cerfolio RJ. The clinical stage of non-small cell lung cancer as assessed by means of fluorodeoxyglucose-positron emission tomographic/computed tomographic scanning is less accurate in cigarette smokers. *The Journal of Thoracic and Cardiovascular Surgery* 2006;**132**(6):1363–8. [PUBMED: 17140957]

Carrillo 2012 *{published data only}*

Carrillo SA, Daniel, VC, Hall N, Hitchcock CL, Ross P Jr, Kassis ES. Fusion Positron emission/Computed tomography underestimates the presence of hilar nodal metastases in patients with resected non-small cell lung cancer. *The Annals of Thoracic Surgery* 2012;**93**(5):1621–4. [PUBMED: 22429676]

Cerfolio 2004 *{published data only}*

Cerfolio RJ, Ojha B, Bryant AS, Raghuveer V, Mountz JM, Bartolucci AA. The accuracy of integrated PET-CT compared with dedicated PET alone for the staging of patients with nonsmall cell lung cancer. *The Annals of Thoracic Surgery* 2004;**78**(3):1017–23. [PUBMED: 15337041]

Cerfolio 2006 *{published data only}*

Cerfolio, RJ, Bryant AS, Eloubeidi MA. Routine mediastinoscopy and esophageal ultrasound fine-needle aspiration in patients with non-small cell lung cancer who are clinically N2 negative: aprospective study. *Chest* 2006; **130**(6):1791–95. [PUBMED: 17166998]

Cerfolio 2007 *{published data only}*

Cerfolio RJ, Bryant AS. Ratio of the maximum standardized uptake value on FDG-PET of the mediastinal (N2) lymph nodes to the primary tumor may be a universal predictor of nodal malignancy in patients with nonsmall-cell lung cancer. *The Annals of Thoracic Surgery* 2007;**83**(5):1826–9. [PUBMED: 17462407]

Cerfolio 2008 *{published data only}*

Cerfolio RJ, Bryant AS. Is palpation of the nonresected pulmonary lobe(s) required for patients with non-small cell lung cancer? A prospective study. *The Journal of Thoracic and Cardiovascular Surgery* 2008;**135**(2):261–68. [PUBMED: 18242247]

Cetinkaya 2011 *{published data only}*

Cetinkaya E, Seyhan EC, Ozgul A, Gencoglu A, Ozgul G, Cam E, Kamiloglu E. Efficacy of convex probe endobronchial ultrasound (CP-EBUS) assisted transbronchial needle aspiration for mediastinal staging in non-small cell lung cancer cases with mediastinal lymphadenopathy. *Annals of Thoracic and Cardiovascular Surgery* 2011;**17**(3):236–42. [PUBMED: 21697783]

Ceylan 2012 *{published data only}*

Ceylan N, Doğan S, Kocaçelebi K, Sava R, Çakan A, Çağrici U. Contrast enhanced CT versus integrated PET-CT in pre-operative nodal staging of non-small cell lung cancer. *Diagnostic and Interventional Radiology* 2012;**18**(5):435–40. [PUBMED: 22374706]

Chiba 2010 *{published data only}*

Chiba K, Isoda M, Chiba M, Kanematsu T, Eguchi S. Significance of PET/CT in determining actual TNM staging for patients with various lung cancers. *International Surgery* 2010;**95**(3):197–204. [PUBMED: 21066996]

Colville 2013 *{published data only}*

Colville D. Impact of F18-FDG PET/CT in suspected or confirmed early stage non-small cell lung cancer. *Nuclear Medicine Communications* 2013;**Conference**:4.

Cömert 2012 *{published data only}*

Cömert SS, Caglayan B, Fidan A, Salepci B, Dogan C, Demirhan R, Ece D. [A comparison of endobronchial ultrasound-guided transbronchial needle aspiration and integrated positron emission tomography-computed tomography in the diagnosis of malignant mediastinal/hilar lymph nodes]. *Türk Göğüs Kalp Damar Cerrahisi Dergisi [Turkish Journal of Thoracic and Cardiovascular Surgery]* 2012;**20**(4):843–49.

Delgado-Bolton 2010 *{published data only}*

Delgado-Bolton RC. Pre-treatment staging with 18F-FDG PET/CT in non-small cell lung cancer (NSCLC): Comparative analysis of the 6th and 7th edition of the TNM classification. *European journal of nuclear medicine and molecular imaging* 2010;**Conference**(var.pagings):S257.

Duan 2012 *{published data only}*

Duan X-B. Analysis of risk factors for mediastinal lymph nodes metastases in non-small cell lung cancer patients with 18F-FDG PET/CT. *Chinese Journal of Medical Imaging Technology* 2012;**28**(6):1135–9.

Duan X-B. Optimal threshold in diagnosing mediastinal lymph node metastasis of non-small cell lung cancer with maximum standardized uptake value. *Chinese Journal of Medical Imaging Technology* 2012;**28**(5):916–20.

Faber 2011 *{published data only}*

Faber DL, Kremer R, Orlovsky M, Lapidot M, Guralnik L, Kagna O, et al.Does PET-CT based clinical staging of non-small cell lung cancer obviate invasive procedures?. *Lung Cancer* 2011;**Conference**:S33–4.

Flechsig 2012 *{published data only}*

Flechsig P. 3D CT-Histogram analysis enables to distinguish affected and FDG-negative lymph nodes in patients with lung cancer. *European Journal of Nuclear Medicine and Molecular Imaging* 2012;**Conference**:S512–3.

Gómez-Caro 2012 *{published data only}*

Gómez-Caro A, Boada M, Cabañas M, Sanchez M, Arguis P, Lomeña F, et al.False-negative rate after positron emission tomography/computer tomography scan for mediastinal staging in cI stage non-small-cell lung cancer. *European Journal of Cardio-Thoracic Surgery* 2012;**42**(1):93–100. [PUBMED: 22290911]

Gregory 2012 *{published data only}*

Gregory DL, Hicks RJ, Hogg A, Binns DS, Shum PL, Milner A, et al.Effect of PET/CT on management of patients with non-small cell lung cancer: results of a prospective study with 5-year survival data. *Journal of Nuclear Medicine* 2012;**53**(7):1007–15. [PUBMED: 22677701]

Günlüoğlu 2010 *{published data only}*

Günlüoğlu MZ, Melek H, Turna A, Medetoğlu B, Kara HV, Demir A, et al.Is there a need for invasive mediastinal staging in centrally located non-small cell lung cancer?. *Chest* 2010;**138**:664A. [DOI: 10.1378/chest.10138]

Günlüoğlu 2011b *{published data only}*

Günlüoğlu, MZ, Melek H, Demir A, Medetoğlu B, Kara HV, Dinçer Sİ . [The accuracy and cost of mediastinal staging strategies for non-small cell lung cancer]. *Türk Göğüs Kalp Damar Cerrahisi Dergisi [Turkish Journal of Thoracic and Cardiovascular Surgery]* 2011;**19**:397–404.

Halpern 2005 *{published data only}*

Halpern BS, Schiepers C, Weber WA, Crawford TL, Fueger BJ, Phelps ME, et al.Presurgical staging of non-small cell lung cancer: positron emission tomography, integrated positron emission tomography/CT, and software image fusion. *Chest* 2005;**128**(4):2289–97. [PUBMED: 16236886]

Hong 2010 *{published data only}*

Hong SP, Song HC, Chong A, Oh JR, Kim JH, Yoo SU, et al.Diagnostic accuracy of preoperative intrathoracic lymph node staging on 18F-FDG PET/CT in patients with non-small cell lung cancer. *Molecular Imaging and Biology* 2010; **Conference**:S1624.

Hu 2008 *{published data only}*

Hu M, Yu JM, Liu NB, Liu LP, Guo HB, Yang GR, et al.[Significance of dual-time-point 18F-FDG PET imaging in evaluation of hilar and mediastinal lymph node metastasis in non-small-cell lung cancer]. *Zhonghua Zhong Liu Za Zhi [Chinese Journal of Oncology]* 2008;**30**(4):306–9. [PUBMED: 18788639]

Huang 2012 *{published data only}*

Huang TW, Hsieh CM, Chang H, Cheng YL, Tzao C, Huang WS. Standard uptake value of positron emission tomography in clinical stage I lung cancer: clinical application and pathological correlation. *European*

Journal of Cardio-Thoracic Surgery 2012;**41**(4):869–73. [PUBMED: 22219418]

Kasai 2010 ***{published data only}***

Kasai T, Motoori K, Horikoshi T, Uchiyama K, Yasufuku K, Takiguchi Y, et al.Dual-time point scanning of integrated FDG PET/CT for the evaluation of mediastinal and hilar lymph nodes in non-small cell lung cancer diagnosed as operable by contrast-enhanced CT. *European Journal of Radiology* 2010;**75**(2):143–6. [PUBMED: 19446975]

Kelly 2006 ***{published data only}***

Kelly A, Cachin F, de Freitas D, Geissler B, Bapt A, Karidioula I, et al.PET-CT in non-small-cell-lung cancer: Clermontoise experience. *Medecine Nucleaire* 2006;**30**(28): 97–106.

Kim 2011 ***{published data only}***

Kim HK, Choi YS, Kim K, Shim YM, Park K, Ahn YC, et al.Outcomes of mediastinoscopy and surgery with or without neoadjuvant therapy in patients with non-small cell lung cancer who are N2 negative on positron emission tomography and computed tomography. *Journal of Thoracic Oncology* 2011;**6**(2):336–342. [PUBMED: 21164366]

Kim 2012 ***{published data only}***

Kim YN, Yi CA, Lee KS, Kwon O, Lee HY, Kim BT, et al.A proposal for combined MRI and PET/CT interpretation criteria for preoperative nodal staging in non-small-cell lung cancer. *European adiology* 2012;**22**(7):1537–46. [PUBMED: 22367469]

Kim 2012a ***{published data only}***

Kim DW, Kim WH, Kim CG. Dual-time-point FDG PET/CT: Is it useful for lymph node staging in patients with non-small-cell lung cancer?. *Nuclear Medicine and Molecular Imaging* 2012;**46**(3):196–200. [PUBMED: 24900060]

Kim 2012b ***{published data only}***

Kim D, Song B, Hong C, Lee S, Ahn B, Lee J. Prediction of occult lymph node metastasis in clinically N0 squamous cell lung carcinoma. *European Journal of Nuclear Medicine and Molecular Imaging* 2012;**Conference**:S568.

Kommata 2011 ***{published data only}***

Kommata S. Clinical value of 18F- FDG-PET/CT scan in nodal staging of patients with non small cell lung cancer. An explorative analysis of retrospective data of 401 patients. *Nuklearmedizin* 2011;**Conference**:6.

Kommata S. Clinical value of 18F-FDG-PET/CT scan in nodal staging of patients with non small cell lung cancer. Evaluation of retrospective data of 401 patients. *European Journal of Nuclear Medicine and Molecular Imaging* 2012; **Conference**:S402–3.

Krueger 2006 ***{published data only}***

Krueger S, Buck A, Pauls S, Halter G, Schumann C, Kropf C, et al.Integrated FDG-PET/CT for improved T and N staging of non-small-cell lung cancer. *European Respiratory Journal* 2006;**28**:S784.

Lapinska 2011 ***{published data only}***

Lapinska GE, Fijolek-Warszewska A, Bryszewska M, Kozlowicz-Gudzinska I, Glogowski M, Zmijewski M, et al.Diagnostic value of PET/CT in preoperative staging in patients with non-small cell lung cancer. *European Journal of Nuclear Medicine and Molecular Imaging* 2011;**Conference**: S359.

Lardinois 2003 ***{published data only}***

Lardinois D, Weder W, Hany TF, Kamel EM, Korom S, Seifert B, et al.Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. *The New England Journal of Medicine* 2003; **348**:2500–7. [PUBMED: 12815135]

Lasnon 2012 ***{published data only}***

Lasnon C. Impact of point spread function reconstruction on pre-therapeutic nodal staging with 18 F -FDG PET in non-small cell lung cancer. *European Journal of Nuclear Medicine and Molecular Imaging* 2010;**Conference**:S318.

Lasnon C, Hicks RJ, Beauregard JM, Milner A, Paciencia M, Guizard AV, et al.Impact of point spread function reconstruction on thoracic lymph node staging with 18F-FDG PET/CT in non-small cell lung cancer. *Clinical Nuclear Medicine* 2012;**37**(10):971–6. [PUBMED: 22899197]

Lebioda 2013 ***{published data only}***

Lebioda A, Makarewicz R, Malkowski B, Dancewicz M, Kowalewski J, Windorbska W. Measurement of primary tumor volume by PET-CT to evaluate risk of mediastinal nodal involvement in NSCLC patients with clinically negative N2 lymph nodes. *Reports of Practical Oncology and Radiotherapy: Journal of Great Poland Cancer Center in Poznan and Polish Society of Radiation Oncology* 2013;**18**(2): 76–81. [PUBMED: 24416539]

Lee 2004 ***{published data only}***

Lee EJ, Choi JY, Choi Y, Choe YS, Lee KH, Lee KS, et al.Improvement in the differentiation of malignant nodes from benign nodes preoperatively in non-small cell lung cancer (NSCLC) using FDG PET/CT. *Journal of Nuclear Medicine* 2004;**45**:374P.

Lee 2008 ***{published data only}***

Lee BE, Redwine J, Foster C, Abella E, Lown T, Lau D, et al.Mediastinoscopy might not be necessary in patients with non-small cell lung cancer with mediastinal lymph nodes having a maximum standardized uptake value of less than 5.3. *The Journal of Thoracic and Cardiovascular Surgery* 2008;**135**(3):615–9. [PUBMED: 18329480]

Lee 2009b ***{published data only}***

Lee HJ, Kim YT, Kang WJ, Lee HJ, Kang CH, Kim JH. Integrated positron-emission tomography for nodal staging in lung cancer. *Asian Cardiovascular and Thoracic Annals* 2009;**17**(6):622–6. [PUBMED: 20026540]

Li 2009 ***{published data only}***

Li M, Wu N, Liang Y, Zheng R, Liu Y, Zhang WJ, et al.[Value of (18)F-FDG PET-CT in the preoperative N staging of non-small cell lung cancer]. *Zhonghua Zhong Liu Za Zhi [Chinese Journal of Oncology]* 2009;**31**(4):288–92. [PUBMED: 19615286]

Li 2011 *{published data only}*

Li X. Mediastinal lymph nodes staging by 18F-FDG PET/CT for early-stage non-small cell lung cancer: A multicenter study. *International Journal of Radiation Oncology Biology Physics* 2011;**Conference**:2.

Li X. Mediastinal lymph nodes staging by 18F-FDG PET/CT for early-stage non-small cell lung cancer: A multicenter study. *Journal of Clinical Oncology* 2011;**Conference**:15.

Li 2012b *{published data only}*

Li X, Zhang H, Xing L, Ma H, Xie P, Zhang L, et al.Mediastinal lymph nodes staging by 18F--FDG PET/CT for early stage non-small cell lung cancer: a multicenter study. *Radiotherapy and Oncology* 2012;**102**(2):246–50. [: 22100657]

Li 2012c *{published data only}*

Li SB, He JX, Li SY, Chen HZ, Yin WQ, Cheng XL, et al.[Value of endobronchial ultrasound-transbronchial needle aspiration biopsy for diagnosis of PET-CT positive mediastinal lymph nodes]. *Zhonghua Zhong Liu Za Zhi [Chinese Journal of Oncology]* 2012;**34**(8):613–5. [PUBMED: 23158997]

Lin 2012 *{published data only}*

Lin WY, Hsu WH, Lin KH, Wang SJ. Role of preoperative PET-CT in assessing mediastinal and hilar lymph node status in early stage lung cancer. *Journal of the Chinese Medical Association* 2012;**75**(5):203–08. [PUBMED: 22632985]

Liu 2009 *{published data only}*

Liu BJ, Dong JC, Xu CQ, Zuo CT, Le JJ, Guan YH, et al.Accuracy of 18F-FDG PET/CT for lymph node staging in non-small-cell lung cancers. *Chinese Medical Journal* 2009;**122**(15):1749–54. [PUBMED: 19781319]

Low 2006 *{published data only}*

Low SY, Eng P, Keng GH, Ng DC. Positron emission tomography with CT in the evaluation of non-small cell lung cancer in populations with a high prevalence of tuberculosis. *Respirology* 2006;**11**(1):84–9. [PUBMED: 16423207]

Ma 2011 *{published data only}*

Ma W, Xu W, Zhu X, Dai D, Song X, Zhu L, et al.Differential diagnosis of mediastinal lymph node metastasis in lung cancer patients using 18F-FDG PET/CT with double quantitative SUVmax and CT value. *Chinese Journal of Clinical Oncology* 2011;**38**:284–7.

Ma W, Xu W, Zhu X, Wang J, Dai D, Song X, et al.Diagnostic value of qualitative and quantitative analysis by 18F-FDG PET/CT for mediastinal lymph nodes of lung cancer. *Chinese Journal of Clinical Oncology* 2011;**38**:512–5.

Maeda 2009 *{published data only}*

Maeda R, Isowa N, Onuma H, Miura H, Harada T, Touge H, et al.The maximum standardized 18F-fluorodeoxyglucose uptake on positron emission tomography predicts lymph node metastasis and invasiveness in clinical stage IA non-small cell lung cancer. *Interactive Cardiovascular and Thoracic Surgery* 2009;**9**(1):79–82. [PUBMED: 19366724]

Mariam 2009 *{published data only}*

Mariam J, Peng C, Clarke J, Thompson JC. Pre-operative radiological staging (CT and PET-CT) compared with pathological staging in patients with resected non-small cell lung cancer attending a regional thoracic centre. *European Journal of Cancer* 2009;**Conference**:S518.

Meduoye 2009 *{published data only}*

Meduoye A, Black E, Duffy J, Beggs D, Majewski A. A comparison of CT/FDG-PET scan staging with final pathological staging in patients following lung resection. *Interactive Cardiovascular and Thoracic Surgery* 2009; **Conference**:S41.

Mendez 2012 *{published data only}*

Mendez M, Maximiano C, Huelves M, Doger de Speville DG, Ibeas P, Ruiz-Valdepeñas A, et al.Mediastinal lymph node staging in patients with non-small cell lung cancer (NSCLC) with F18-fluorodeoxyglucose positron emission tomography/computed tomography (FDG-PET/CT). *Journal of Clinical Oncology* 2012;**Conference**:15.

Mi 2012 *{published data only}*

Mi B, Wan W, Yu C, You X, Jiang F, You Q. [The value of ^{18} F-FDG PET/CT in improving diagnosis of lung cancer].. *Zhongguo Fei Ai Za Zhi [Chinese Journal of Lung Cancer]* 2012;**15**(2):78–83. [PUBMED: 22336234]

Moodie 2009 *{published data only}*

Moodie K, Cherk MH, Lau E, Turlakow A, Skinner S, Hicks R, et al.Evaluation of pulmonary nodules and lung cancer with one-inch crystal gamma coincidence positron emission tomography/CT versus dedicated positron emission tomography/CT. *Journal of Medical Imaging and Radiation Oncology* 2009;**53**(1):32–39. [PUBMED: 19453526]

Moreno Garcia 2009 *{published data only}*

Moreno Garcia V, Castro JD, Feliu J, Belda C, Barriuso J, Marin MD, et al.Accuracy of integrated PET-CT for mediastinal lymph node metastases in non-small cell lung cancer. *European Journal of Cancer* 2009;**Conference**:S169.

Morikawa 2011 *{published data only}*

Morikawa M. Clinical usefulness of positron emission tomography/computed tomography and magnetic resonance imaging for lung cancer. *American Journal of Respiratory and Critical Care Medicine* 2011;**Conference**:1.

Nakajima 2011 *{published data only}*

Nakajima E, Sakata Y, Saji H, Kato Y, Suga Y, Uchino Y, et al.The consensus to evaluate mediastinal lymph node metastasis in primary lung cancer with integrated FDG PET/CT. *American Journal of Respiratory and Critical Care Medicine* 2011;**Conference**:1.

Nakamura 2008 *{published data only}*

Nakamura H, Taguchi M, Kitamura H, Nishikawa J. Fluorodeoxyglucose positron emission tomography integrated with computed tomography to determine resectability of primary lung cancer. *General Thoracic and*

Cardiovascular Surgery 2008;**56**(8):404–9. [PUBMED: 18696206]

Nomori 2008 *{published data only}*
Nomori H, Mori T, Ikeda K, Kawanaka K, Shiraishi S, Katahira K, et al.Diffusion-weighted magnetic resonance imaging can be used in place of positron emission tomography for N staging of non-small cell lung cancer with fewer false-positive results. *The Journal of Thoracic and Cardiovascular Surgery* 2008;**135**(4):816–22. [PUBMED: 18374761]

Ohno 2007 *{published data only}*
Ohno Y, Koyama H, Nogami M, Takenaka D, Yoshikawa T, Yoshimura M, et al.STIR turbo SE MR imaging vs. coregistered FDG-PET/CT: quantitative and qualitative assessment of N-stage in non-small-cell lung cancer patients. *Journal of Magnetic Resonance Imaging* 2007;**26**(4):1071–80. [PUBMED: 17896365]

Ozkan 2010 *{published data only}*
Ozkan ZG. How efficient is PET/CT in metabolic characterization of lung lesions?. *European Journal of Nuclear Medicine and Molecular Imaging* 2010;**Conference**: S423.

Özta 2012 *{published data only}*
Özta S, Öztürk AV, Acartürk E, Tezel Y, Özdemir M, Ataç G, et al.The role of tumor SUVmax/lymph node SUVmax ratio viewed on PET-CT in the detection of mediastinal metastasis in patients with lung cancer. *Türk Göğüs Kalp Damar Cerrahisi Dergisi [Turkish Journal of Thoracic and Cardiovascular Surgery]* 2012;**20**(3):544–51.

Pauls 2012 *{published data only}*
Pauls S, Schmidt SA, Juchems MS, Klass O, Luster M, Reske SN, et al.Diffusion-weighted MR imaging in comparison to integrated [^{18} F]-FDG PET/CT for N-staging in patients with lung cancer. *European Journal of Radiology* 2012;**81**(1):178–82. [PUBMED: 20932700]

Pfannenberg 2007 *{published data only}*
Pfannenberg AC, Aschoff P, Brechtel K, Müller M, Bares R, Paulsen F, et al.Low dose non-enhanced CT versus standard dose contrast-enhanced CT in combined PET/CT protocols for staging and therapy planning in non-small cell lung cancer. *European Journal of Nuclear Medicine and Molecular Imaging* 2007;**34**(1):36–44. [PUBMED: 16896664]

Pozo-Rodriguez 2005 *{published data only}*
Pozo-Rodriguez F, Martin de Nicolás JL, Sánchez-Nistal MA, Maldonado A, Garcia de Barajas S, Calero-Garcia R, et al.Accuracy of helical computed tomography and [18F] fluorodeoxyglucose positron emission tomography for identifying lymph node mediastinal metastases in potentially resectable non-small-cell lung cancer. *Journal of Clinical Oncology* 2005;**23**(33):8348–56. [PUBMED: 16219937]

Prévost 2009 *{published data only}*
Prévost A, Papathanassiou D, Jovenin N, Menéroux B, Cuif-Job A, Bruna-Muraille C, et al.[Comparison between PET (-FDG) and computed tomography in the staging of lung cancer. Consequences for operability in 94 patients]. *Revue de Pneumologie Clinique* 2009;**65**(6):341–9. [PUBMED: 19995654]

Quaia 2008 *{published data only}*
Quaia E, Tona G, Gelain F, Lubin E, Pizzolato R, Boscolo E, et al.Integrated fluorine-18 fluorodeoxyglucose (18F-FDG) PET/CT compared to standard contrast-enhanced CT for characterization and staging of pulmonary tumors eligible for surgical resection. *Acta Radiologica* 2008;**49**(9):995–1004. [PUBMED: 18651256]

Salman 2009 *{published data only}*
Salman KA, Steinmann CH, Sukumar VP, Schulthess GK, Steinert HC. PET/CT staging of T1-stage non-small cell lung cancer. *Internal Medicine Journal* 2009;**39**(Suppl4):A110.

Sánchez Sánchez 2011 *{published data only}*
Sánchez Sánchez R, Rodríguez Fernández A, Gómez Río M, Alkurdi Martínez A, Castellón Rubioc VE, Ramos Font C, et al.[Utility of PET/CT for mediastinal staging of non-small cell lung cancer in stage III (N2)]. *Revista Espanola de Medicina Nuclear* 2011;**30**(4):211–6. [PUBMED: 21514978]

Schiavariello 2012 *{published data only}*
Schiavariello S. Lymph-node mediastinal staging in patients with NSCLC: 18-FDG PET-TC versus TC MDC. *European Journal of Nuclear Medicine and Molecular Imaging* 2012;**Conference**:S454.

Schreyögg 2010 *{published data only}*
Schreyögg J, Weller J, Stargardt T, Herrmann K, Bluemel C, Dechow T, et al.Cost-effectiveness of hybrid PET/CT for staging of non-small cell lung cancer. *Journal of Nuclear Medicine* 2010;**51**(11):1668-75. [PUBMED: 21051648]

Schwenzer 2012 *{published data only}*
Schwenzer NF, Schraml C, Mueller M, Brendle C, Sauter A, Spengler W, et al.Pulmonary lesion assessment: comparison of whole-body hybrid MR/PET and PET/CT imaging--pilot study. *Radiology* 2012;**264**(2):551–58. [PUBMED: 22653189]

Shim 2005 *{published data only}*
Shim SS, Lee KS, Kim BT, Chung MJ, Lee EJ, Han J, et al.Non-small cell lung cancer: prospective comparison of integrated FDG PET/CT and CT alone for preoperative staging. *Radiology* 2005;**236**(3):1011–19. [PUBMED: 16014441]

Sit 2010 *{published data only}*
Sit AK, Sihoe AD, Suen WS, Cheng LC. Positron-emission tomography for lung cancer in a tuberculosis-endemic region. *Asian Cardiovascular and Thoracic Annals* 2010;**18**(1):33–8. [PUBMED: 20124294]

Sivrikoz 2010 *{published data only}*
Sivrikoz MC. Mediastinoscopy and F-18 PET-CT in NSCLC. *European Journal of Nuclear Medicine and Molecular Imaging* 2010;**Conference**(var.pagings):S318.

Sivrikoz 2012 *{published data only}*
Sivrikoz C, Ak I, Simsek FS, Döner E, Dündar E. Is Mediastinoscopy Still the Gold Standard to Evaluate Mediastinal Lymph Nodes in Patients with Non-Small Cell Lung Carcinoma?. *The Thoracic and Cardiovascular Surgeon* 2012;**60**(2):116–21. [PUBMED: 21692019]

Steinert 2010 *{published data only}*
Steinert H. Outcome of PET/CT imaging in staging of non-small-cell lung cancer <= 3cm in diameter. *Nuklearmedizin* 2010;**Conference**:2.

Tamura 2012 *{published data only}*
Tamura, M, Oda M, Matsumoto I, Waseda R, Watanabe G. Pattern and predictors of false positive lymph node involvement on positron emission tomography in patients with non-small cell lung Cancer. *The Thoracic and Cardiovascular Surgeon* 2012;**60**(2):105–10. [PUBMED: 21789758]

Tsutani 2012 *{published data only}*
Tsutani Y, Miyata Y, Nakayama H, Okumura S, Adachi S, Yoshimura M, et al.Prediction of pathologic node-negative clinical stage IA lung adenocarcinoma for optimal candidates undergoing sublobar resection. *The Journal of Thoracic and Cardiovascular Surgery* 2012;**144**(6):1365–71. [PUBMED: 22883546]

Vaz 2012 *{published data only}*
Vaz AP, Fernandes G, Souto Moura C, Bastos P, Queiroga H, Hespanhol V. Integrated PET/CT in non-small cell lung cancer staging--clinical and pathological agreement. *Revista Portuguesa de Pneumologia* 2012;**18**(3):109–114. [PUBMED: 22405953]

Ventura 2010 *{published data only}*
Ventura E, Islam T, Gee MS, Mahmood U, Braschi M, Harisinghani MG. Detection of nodal metastatic disease in patients with non-small cell lung cancer: comparison of positron emission tomography (PET), contrast-enhanced computed tomography (CT), and combined PET-CT. *Clinical Imaging* 2010;**34**(1):20–8. [PUBMED: 20122515]

Wang 2012 *{published data only}*
Wang F, Ma S, Shen L, Li N, Yang Z, Chen K. [Application of ^{18} F-FDG PET/CT in pulmonary disease: a report of 419 cases]. *Zhongguo Fei Ai Za Zhi [Chinese Journal of Lung Cancer]* 2012;**15**(1):21–6. [PUBMED: 22237120]

Wiese 2012 *{published data only}*
Wiese MN. Preoperative staging of non-small-cell lung cancer: Comparison of whole body diffusion weighted magnetic resonance imaging and 18f-fluorodeoxyglucose positron emission tomography/computed tomography. *Interactive Cardiovascular and Thoracic Surgery* 2012; **Conference**:S57–8.

Wu 2010 *{published data only}*
Wu CS. Preoperative lymph node staging in patients of non-small-cell lung cancer: accuracy of 18F-FDG PET/CT. *European Journal of Nuclear Medicine and Molecular Imaging* 2010;**Conference**:S424.

Yi 2007 *{published data only}*
Yi CA, Lee KS, Kim BT, Shim SS, Chung MJ, Sung YM, et al.Efficacy of helical dynamic CT versus integrated PET/CT for detection of mediastinal nodal metastasis in non-small cell lung cancer. *American Journal of Roentgenology* 2007; **188**(2):318–25. [PUBMED: 17242237]

Yi 2011 *{published data only}*
Yi CA, Lee KS, Kim YN, Lee E, Kwon OJ, Kim BT, et al.Preoperative nodal staging in patients with non-small cell lung cancer: Comparison between multimodality MR imaging plus PET/CT and PET/CT alone. *Journal of Thoracic Imaging* 2011;**26**:W106.

Yi 2013 *{published data only}*
Yi CA, Lee KS, Lee HY, Kim S, Kwon OJ, Kim H, et al.Coregistered whole body magnetic resonance imaging-positron emission tomography (MRI-PET) versus PET-computed tomography plus brain MRI in staging resectable lung cancer: Comparisons of clinical effectiveness in a randomized trial. *Cancer* 2013;**119**(10):1784-91. [PUBMED: 23423920]

Yu 2007 *{published data only}*
Yu L-J, Duan Y, Liang X-Y, Wang X. 18FDG PET/CT in diagnosis and metastasis detection of lung neoplasms. *Chinese Journal of Medical Imaging Technology* 2007;**23**: 605–7.

Zhang 2006 *{published data only}*
Zhang X, Li TR, Chen ZS, Ouyang XN. [Comparing serum tumor antigen detection combined with CT scan with PET-CT for lung cancer diagnosis]. *Ai Zheng [Chinese Journal of Cancer]* 2006;**25**(1):66–8. [PUBMED: 16405752]

Zhang 2012 *{published data only}*
Zhang TM, Zhang LM, Liu Y, Zhang ZF, Wang CL. [Diagnostic value of (18)F-FDG PET/CT imaging plus serum tumor marker assays for pulmonary lesions and clinical significance of SUVmax]. *Zhonghua Yie Xue Za Zhi* 2012;**92**(41):2901–4. [PUBMED: 23328236]

Zsiray 2009 *{published data only}*
Zsiray M, Markoczy Z, Magyar M, Lengyel Z, Fekeshazy A, Borbely L. [The advantage of positron emission tomography combined with computer tomography (PET-CT) in the diagnosis of lung cancer, experience with 408 patients]. *Magyar Onkologia* 2009;**53**(1):17–21. [PUBMED: 19318322]

Zsiray 2011 *{published data only}*
Zsiray M, Markoczy Z, Lengyel Z, Fekeshazy A, Kasler M, Borbely K. Contribution of 18F-FDG PET/CT in the management of patients with lung cancer. *European Journal of Nuclear Medicine and Molecular Imaging* 2011; **Conference**:S125–5.

Zsiray 2012 *{published data only}*
Zsiray M. Preoperative mediastinal staging of NSCLC: the clinical impact of 18F-FDG PET/CT. *European Journal of Nuclear Medicine and Molecular Imaging* 2012;**Conference**: S455–6.

Additional references

Casali 2010
Casali C, Cucca M, Rossi G, Barbieri F, Iacuzio L, Bagni B, et al.The variation of prognostic significance of maximum standarized uptake value of [18F]-fluoro-2-deoxy-glucose positron emission tomography in different histological subtypes and pathological stages of surgically resected non-small cell lung carcinoma. *Lung Cancer* 2010;**69**(2): 187–93. [PUBMED: 19942313]

Cerfolio 2005
Cerfolio RJ, Bryant AS, Ojha B, Eloubeidi M. Improving the inaccuracies of clinical staging of patients with NSCLC: a prospective trial. *The Annals of Thoracic Surgery* 2005;**80** (4):1207–13. [PUBMED: 16181842]

Davidson 2009
Davidson JA, Wong V, Fraser R, Hirsch V. Comparison of primary tumor maximal standardized uptake value (SUVmax) on preoperative [18F]fluorodeoxyglucose positron emission tomography/computed tomography (PET/CT) and histological subtype in patients with non-small cell lung cancer (NSCLC). *Journal of Clinical Oncology* 2009;**27**(15s):abstr7571.

De Leyn 2014
De Leyn P, Dooms C, Kuzdzal J, Lardinois D, Passlick B, Rami-Porta R, et al.Revised ESTS guidelines for preoperative mediastinal lymph node staging for non-small-cell lung cancer. *European Journal of Cardiothoracic Surgery* 2014;**45**(5):787–98. [PUBMED: 24578407]

Department of Health 2011
Department of Health. Improving Outcomes: a strategy for cancer. http://www.dh.gov.uk/prod˙consum˙dh/groups/dh˙digitalassets/documents/digitalasset/dh˙123394.pdf 2011.

Detterbeck 2007
Detterbeck FC, Jantz MA, Wallace M, Vansteenkiste J, Silvestri GA. Invasive mediastinal staging of lung cancer. *Chest* 2007;**132 Suppl**:202S–220S. [DOI: 10.1378/chest.07-1362]

Harbord 2007
Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;**8**(2):239–51. [PUBMED: 16698768]

Jeong 2002
Jeong HJ, Min JJ, Park JM, Chung JK, Kim BT, Jeong JM, et al.Determination of the prognostic value of [18F]fluorodeoxyglucose uptake by using positron emission tomography in patients with non-small cell lung cancer. *Nuclear Medicine Communications* 2002;**23**(9):865–70. [PUBMED: 12195091]

Kim 2006
Kim B-T, Lee KS, Shim SS, Choi JY, Kwon OJ, Kim H, et al.Stage T1 non-small cell lung cancer: preoperative mediastinal nodal staging with integrated FDG PET/CT -- a prospective study. *Radiology* 2006;**241**(2):501–9. [PUBMED: 16966480]

Lee 2007a
Lee BE, von Haag D, Lown T, Lau D, Calhoun R, Follette D. Advances in positron emission tomography technology have increased the need for surgical staging in non-small cell lung cancer. *The Journal of Thoracic and Cardiovascular Surgery* 2007;**133**(3):746–52. [PUBMED: 17320577]

Lindholm 2014
Lindholm H, Staaf J, Jacobsson H, Brolin F, Hatherly R, Sânchez-Crespo A. Repeatability of the Maximum Standard Uptake Value (SUVmax) in FDG PET. *Molecular Imaging and Radionuclide Therapy* 2014;**23**(1):16–20. [PUBMED: 24653930]

Lu 2010
Lu P, Yu L, Li Y, Sun Y. A correlation study between maximum standardized uptake values and pathology and clinical staging in nonsmall cell lung cancer. *Nuclear Medicine Communications* 2010;**31**(7):646–51. [PUBMED: 20545045]

Macaskill 2004
Macaskill P. Empirical Bayes estimates generated in a hierarchical summary ROC analysis agreed closely with those of a full Bayesian analysis. *Journal of Clinical Epidemiology* 2004;**57**:925-32..

Maemondo 2010
Maemondo M, Inoue A, Kobayashi K, Sugawara S, Oizumi S, Isobe H, et al.Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *The New England Journal of Medicine* 2010;**362**(25):2380–8. [PUBMED: 20573926]

Manser 2005
Manser R, Wright G, Hart D, Byrnes G, Campbell D, Wainer Z, et al.Surgery for local and locally advanced non-small cell lung cancer. *Cochrane Database of Systematic Reviews* 2005, Issue 1. [DOI: 10.1002/14651858.CD004699.pub2]

NICE 2005
National Institute for Health and Clinical Excellence (NICE). Referral guidelines for suspected cancer. http://guidance.nice.org.uk/CG27/Guidance 2005.

NICE 2011
National Institute for Health and Clinical Excellence (NICE). The diagnosis and treatment of lung cancer (update). http://guidance.nice.org.uk/CG121 2011.

O'Rourke 2010
O'Rourke N, Roqué i Figuls M, Farré Bernadó N, Macbeth F. Concurrent chemoradiotherapy in non-small cell lung cancer. *Cochrane Database of Systematic Reviews* 2010, Issue 6. [DOI: 10.1002/14651858.CD002140.pub3]

Reitsma 2005
Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005; **58**(10):982–90. [PUBMED: 16168343]

Review Manager 2012

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012.

Rutter 2001

Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in Medicine* 2001;**20**(19):2865–84. [PUBMED: 11568945]

Shim 2005

Shim SS, Lee KS, Kim BT, Chung MJ, Lee EJ, Han J, et al.Non-small cell lung cancer: prospective comparison of integrated FDG PET/CT and CT alone for preoperative staging. *Radiology* 2005;**236**(3):1011–9. [PUBMED: 16014441]

Silvestri 2013

Silvestri GA, Gonzalez AV, Jantz MA, Margolis ML, Gould MK, Tanoue LT, et al.Methods for staging non-small cell lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest* 2013;**143**(5 Suppl): e211S–50S. [PUBMED: 23649440]

Tournoy 2007

Tournoy KG, Maddens S, Gosselin R, Van Maele G, van Meerbeeck JP, Kelles A. Integrated FDG-PET/CT does not make invasive staging of the intrathoracic lymph nodes in non-small cell lung cancer redundant: a prospective study. *Thorax* 2007;**62**(8):696–701. [PUBMED: 17687098]

Whiting 2011

Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al (the QUADAS-2 Group). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8): 529–36. [PUBMED: 22007046]

Yi 2007

Yi CA, Lee KS, Kim BT, Shim SS, Chung MJ, Sung YM, et al.Efficacy of helical dynamic CT versus integrated PET/CT for detection of mediastinal nodal metastasis in non-small cell lung cancer. *American Journal of Roentgenology* 2007; **188**(2):318–25. [PUBMED: 17242237]

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Bille 2013

Study characteristics	
Patient sampling	Retrospective consecutive patient series
Patient characteristics and setting	353 participants, median age = 68 (range = 37 to 86) years, 258 males/95 females, UK (?) Italy **Histology of primary tumour** Adenocarcinoma: N = 244; squamous cell: N = 109; comorbidities: not reported **Inclusion criteria** Consecutive participants who underwent surgery (either mediastinoscopy, anterior mediastinotomy, thoracotomy, or a combination of the aforementioned) for suspected or pathologically proven localised, clinically resectable NSCLC over a 6-year period between August 2004 and January 2010 **Exclusion criteria** Participants who had PET-CT performed elsewhere, who received induction chemotherapy, radiation therapy, or both; those with a PET-CT-negative primary tumour; and those with histological types other than adenocarcinoma and squamous cell carcinoma **Previous tests** Conventional diagnostic work-up, including a thorough history and physical examination; laboratory tests; spirometry; chest X-ray; contrast-enhanced brain, chest, and upper abdomen CT; and bronchoscopy **Clinical setting** Thoracic surgery unit The inclusion only of participants who received surgery for adeno- and squamous cell NSCLC narrows the range of patients who would receive the index test in standard practice
Index tests	Participants were asked to fast for at least 6 h before the examination, and a serum glucose level below 160 mg/dl was ensured. Image acquisition using an integrated PET-CT scanner (Discovery ST; GE Medical systems) was performed 60 min after intravenous administration of FDG (3.5 to 4.5 MBq/kg). The CT scan was used for both anatomical localisation and for the calculation of attenuation correction. The integrated PET-CT data sets were prospectively evaluated in consensus by 2 nuclear medicine physicians (EP and VA) who were aware of the clinical and stand-alone contrast-enhanced CT results, but blinded to the histological findings. Pulmonary and mediastinal lymph node stations were deemed positive for metastatic spread if they exhibited focally increased FDG uptake higher than the normal background activity, as determined by qualitative analysis **Covariates** Type of PET-CT scanner: integrated PET/CT scanner (Discovery ST; GE Medical systems) FDG dose: 3.5 to 4.5 MBq/kg Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): focally increased FDG uptake higher than the normal background activity. There was no prespecified cut-off value
Target condition and reference standard(s)	Surgical staging (either mediastinoscopy, anterior mediastinotomy, thoracotomy, or a combination of the aforementioned). Invasive mediastinal staging procedures were performed in participants (n = 41) considered N2/N3 lymph node positive by PET/CT. Cervical mediastinoscopy was used to

	sample stations 2R, 4R, 2L, 4L and 7, and anterior mediastinotomy was used to sample stations 5 and 6
Flow and timing	The paper reports that 64/413 participants were excluded due to NSCLC other than adenocarcinoma or squamous cell carcinoma. This leaves 349 participants, not 353 as was reported. Integrated PET-CT was performed no more than 3 weeks prior to surgery
Comparative	
Notes	There was no mention of funding source, but since this was a retrospective database study, it is likely that the study received no explicit funding Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Bryant 2006a

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	397 participants, median age = 67 (range = 23 to 82) years, 251 males/146 females, US **Histology of primary tumour** Not reported; comorbidities: not reported **Inclusion criteria** Participants who presented to 1 surgeon with an indeterminate pulmonary nodule or biopsy-proven NSCLC and who underwent integrated FDG PET-CT at the authors' institution from January 2003 to December 2004 **Exclusion criteria** Participants < 19 years old with type I diabetes, Pancoast tumour, T4 tumour from mediastinal invasion or from malignant effusion, neoadjuvant chemotherapy or radiation, or biopsy-proven metastatic (M1) disease **Previous tests** Participants were clinically TNM-staged based on the FDG PET-CT results **Clinical setting** Cardiothoracic surgery unit
Index tests	The FDG PET-CT scans were performed on an integrated FDG PET-CT scanner (GE Discovery LS PET-CT scanner; Milwaukee, WI, USA). Participants were asked to fast for 4 hours and then subsequently received 555 MBq (15 mCi) of FDG intravenously followed by PET after 1 hour. The scans were performed from the skull base to mid-thigh level. The CT examination was used for attenuation correction of PET images. The scanning time for emission PET was 5 minutes

	per bed position. The most recent CT scan of the chest was also available for visual correlation. Maximum SUV (maxSUV) of the primary lung lesion and of each suspicious lymph node station was determined by drawing regions of interest on the attenuation-corrected FDG-PET images around it. The maxSUV within the selected region of interest was used throughout the study exclusively, and N2, N3, or M1 areas with maxSUV $\geq$ 2.5 were considered suspicious **Covariates** Type of PET-CT scanner: integrated FDG-PET-CT scanner (GE Discovery LS PET-CT Scanner; Milwaukee, WI, USA) FDG dose: 555 MBq (15 mCi) Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): maxSUV $\geq$ 2.5
Target condition and reference standard(s)	All suspicious N2, N3, or M1 areas (maxSUV $\geq$ 2.5) were biopsied prior to pulmonary resection. Mediastinoscopy was used to biopsy suspicious lymph nodes in the paratracheal area (stations 2R, 4R, 2L, and 4L) and the superior portion of the subcarinal lymph node. Either video-assisted thoracoscopic surgery (VATS) or endoscopic transoesophageal ultrasound (EUS) was used to biopsy suspicious posterior aortopulmonary window lymph nodes, subcarinal lymph nodes, perioesophageal nodes, and inferior pulmonary ligament nodes. Microscopic disease was defined as tumour invasion of $\leq$ 2 mm or disease only detected on immunohistochemical staining. The latter was performed in selected cases only as per the pathologist's discretion. In general, there were 2 slices per lymph nodes used
Flow and timing	All participants were accounted for in the results. Apart from 26/397 participants whose data were not included (because of proven M1 disease/T4 tumour/Pancoast tumour or refusal of definitive surgery), all the remaining participants received verification using the reference standard
Comparative	
Notes	26/397 participants had proven M1 disease/T4 tumour/Pancoast tumour or refused definitive surgery and their data were not included The paper makes no mention of potential sources of support. The data appear to have been gathered as part of normal practice Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Carnochan 2009

Study characteristics			
Patient sampling	Retrospective consecutive (?) patient series		
Patient characteristics and setting	200 participants, 95 females (mean age = 63.7 years, SD = 9.2 years, range = 40 to 81 years) and 105 males (mean age = 66.4 years, SD = 7.9 years, range = 44 to 84 years), Scotland **Histology of primary tumour** Not reported; comorbidities: not reported **Inclusion criteria** Participants referred for consideration of surgery with a diagnosis or presumptive diagnosis of bronchogenic carcinoma to the authors' unit between June 2006 and January 2008 **Exclusion criteria** None listed **Previous tests** Spiral CT (all participants), and additional imaging (including ultrasound, MRI, and isotope scanning) was utilised where appropriate **Clinical setting** Thoracic surgery unit		
Index tests	The participants were referred from different regions, which meant that both the primary CT and subsequent PET-CT imaging was undertaken at different sites. All PET-CT scans were reported by 2 radiologists, with node positivity based on SUV criteria selected by the scanning unit **Covariates** Type of PET-CT scanner: not reported, but undertaken at different sites FDG dose: not reported Injection-to-scan time: not reported Attenuation correction: not reported Cut-off values for test positivity (malignancy): node positivity based on SUV criteria selected by the scanning unit		
Target condition and reference standard(s)	Mediastinoscopy with or without surgical resection (performed within 4 weeks of the PET-CT scan)		
Flow and timing	Data from 194/200 were available. The remaining 6 participants were classified as having benign disease by PET-CT, but histology later showed that they all had NSCLC. However, the N-stage for these participants was not reported		
Comparative			
Notes	There was no mention of funding source, but since this was a retrospective database study, it is likely that the study received no explicit funding Adverse events: not reported		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	No		
			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Chen 2010

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	56 participants, mean age = 51 (range = 35 to 76) years, 35 males/21 females, China **Histology of primary tumour** Not reported. comorbidities: not reported **Inclusion criteria** Quote: "From January to March in 2008, 56 consecutive participants with NSCLC underwent WB-DWI and integrated FDG PET/CT for primary tumor, lymph node metastasis and distant metastasis." No further information reported **Exclusion criteria** Not reported **Previous/all reported tests** See **Inclusion criteria**. No further information reported **Clinical setting** Department of Radiology
Index tests	All participants fasted for at least 4 to 6 hours, and normal blood glucose levels were verified on blood samples collected before intravenous administration of FDG at a rate of 3.3 MBq/kg. PET-CT images were obtained from the skull to the mid thigh 60 min after completion of injection. PET-CT imaging was conducted by using the Biograph mMR system, which consisted of a PET scanner (Siemens Medical Solutions, Hoffman Estates, IL, USA) and a 2-section CT scanner (Siemens Medical Solutions, Erlangen, Germany). PET data were collected with a full-ring PET tomography. The PET component of the Biograph mMR had an in-plane spatial resolution of 4.6 mm and a transverse field of view of 15.5 cm for 1 table position. After unenhanced CT was performed, emission PET was performed in the identical transverse field of view. The system generated separate CT and PET data sets that could be fused by using a syngo-based fusion tool (Siemens Medical Solutions, Erlangen, Germany). The whole PET/CT study took approximately 45 min. All FDG-PET-CT studies were independently reviewed by 2 nuclear medicine physicians, who were also blinded to all information about the results of whole-body MR and conventional radiologic examinations **Covariates** Type of PET-CT scanner: Biograph system (Siemens). FDG dose: 3.3 MBq/kg Injection-to-scan time: 60 min Attenuation correction: probably, but not explicitly reported Cut-off values for test positivity (malignancy): (quote (author communication)) "The criteria for test positivity on PET-CT is visually more metabolically active than mediastinal blood pool"
Target condition and reference standard(s)	All primary tumours were diagnosed on the basis of pathologic results from endoscopic CT-guided or surgical biopsies. The metastastic sites were determined on the basis of the results of CT, integrated FDG PET-CT, and MR examinations. The diagnosis of a lesion as metastasis was determined either by pathology or follow-up. (The lesion became larger during the follow-up periods or decreased in size after treatment.) The follow-up was maintained for more than 6 months in every participant,

	and the diameters of suspect (?) lymph nodes (>10 mm) were consecutively observed; if no change in size was observed during the period, then a diagnosis as nonmetastatic lymph nodes was made
Flow and timing	It is unclear how many participants received pathological confirmation of their N status and how many received follow-up as the reference standard
Comparative	
Notes	Funding: no details reported Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	No		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Chen 2010 (*Continued*)

Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	No		

Czepczynski 2011

Study characteristics	
Patient sampling	Patient series
Patient characteristics and setting	51 participants, median age = not reported (range = 39 to 73 years), numbers of males/females not reported, Poland **Histology of primary tumour** Not reported; comorbidities: not reported **Inclusion criteria** Participants with newly diagnosed stage I to IIIa NSCLC who were treated with surgery at the authors' institutions in 2008 to 2009 **Exclusion criteria** None listed **Previous/all reported tests** Not reported **Clinical setting** Secondary/tertiary setting The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	**Covariates** Type of PET-CT scanner: Discovery ST PET-CT scanner (GE) FDG dose: 5 MBq/kg Injection-to-scan time: 50 to 70 min

Czepczynski 2011 (*Continued*)

	Attenuation correction: not reported Cut-off values for test positivity (malignancy): not reported
Target condition and reference standard(s)	Histopathology after anatomical resection of the lung tumour and the mediastinal lymph nodes ≤ 6 weeks after the PET-CT
Flow and timing	51/440 participants with NSCLC who had PET-CT scans in 2008 to 2009 "were qualified to the study"
Comparative	
Notes	Author sent test accuracy data on request. As the study was only published as an abstract, we contacted the author (on 5 November 2012) to request the following information: How was the sample recruited? - Did the study population consist of a consecutive sample? - Was the recruitment prospective or retrospective? Characteristics of the 51 participants: - Median age = () years - (number of males) males/(number of females) females - Histology of primary tumour? () - Comorbidities: () - Inclusion criteria? - Exclusion criteria? - Previous/all reported tests? Clinical setting: - Thoracic surgery unit? - Were any participants excluded from the analyses? PET-CT scanning: - Did you use attenuation correction? - What was the criteria for a positive result? - Did you use a prespecified cut-off value for test positivity? - Was the PET-CT results interpreted without knowledge of the pathological results? Pathological staging - Was the pathological staging results interpreted without knowledge of the PET-CT results? Funding - Was the study funded and if yes, by whom? We received no response Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	No		
			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		

Darling 2011

Study characteristics	
Patient sampling	Prospective randomised patient series
Patient characteristics and setting	149 participants, median age = 67 (range = 41 to 86) years, 76 males/73 females, Canada **Histology of primary tumour** Adenocarcinoma: N = 64; squamous cell carcinoma: N = 24; large cell carcinoma: N = 5; not otherwise specified (NOS): N = 53; suspicious for NSCLC: N = 3; comorbidities: not reported **Inclusion criteria** Participants from 4 academic metropolitan tertiary centres and 4 community hospitals in Ontario (Canada) with histologically or cytologically proven NSCLC clinical stage I, II, or IIIA disease based on CT chest; and who were considered candidates for surgical resection between 2004 and 2007 **Exclusion criteria** Participants with poor pulmonary function, poor performance status (ECOG grade 3 to 4) or clinically significant concurrent medical problems making them unfit for surgery, unable to lie supine for PET, with cancer (unless they had been disease-free for $\geq$ 5 years, had non-melanotic skin cancer, or had carcinoma in situ of the cervix), or had undergone part of the staging strategy under investigation within 8 weeks of randomisation. Please note, the included participants were only those who were randomised to PET-CT and not those randomised to conventional staging (abdominal CT and bone scan) **Previous/all reported tests** History taking, physical examination, routine blood analysis (including random glucose, creatinine, liver enzyme, and alkaline phosphatase measurement), chest CT, and cranial (MRI or CT) imaging **Clinical setting** Secondary/tertiary setting
Index tests	Participants were imaged after they had fasted for at least 4 hours. Participants were injected with FDG, 5 MBq/kg of body weight (10%), to a maximum of 500 MBq. Blood glucose was measured before injection of FDG in all participants; if the fasting blood glucose level was greater than 9.7 mmol/L (175 mg/dL), the study was delayed until adequate diabetic control had been established. Participants rested quietly for 60 minutes after tracer injection. For participants who had PET-CT, a low-dose CT scan was initially acquired with the axial field of view extending from the base of the skull to mid thigh. Immediately thereafter, emission data acquisitions were obtained over the same field of view (Biograph Duo and Gemini Dual, 3 min/bed; Discovery LS4, 5 min/bed; and ECAT ART (a PET scanner equipped with a partial ring of bismuth germanate detectors), 10 minutes/bed). For participants studied with the ECAT ART, transmission data acquisition of 4 min/bed followed the emission acquisition over the same axial field of view All PET-CT images were interpreted at the site where the PET study was performed. The interpreter's degree of suspicion for an abnormality was recorded by using a 5-point ordinal scale with the following categories: 0 = normal; 1 = probably normal; 2 = equivocal; 3 = probably abnormal; and 4 = definitely abnormal. The physicians who interpreted the PET scans were free to use information from the standardised uptake value determination to assist in grading of the identified abnormalities according to the 5-point scale. The readers were not provided with cut-off specific uptake values to determine the presence or absence of cancer **Covariates** Type of PET-CT scanner: 5 different scanners used: Discovery LS4 (General Electric, Waukesha, Wisconsin; N = 64), Biograph Duo (CTI/Siemens, Knoxville, Tennessee; N = 55), 2 Philips Gemini Dual machines (Philips Electronics NV, Eindhoven, the Netherlands; N = 39 and 7, respectively), and the ECAT ART (CTI/Siemens; N = 4)

	FDG dose: 5 MBq/kg of body weight, to a maximum of 500 MBq Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): not reported/qualitative
Target condition and reference standard(s)	Pathological staging from cervical mediastinoscopy, anterior mediastinotomy, thoracotomy, or a combination of the aforementioned, with detailed lymph node sampling
Flow and timing	All participants received the same reference standard, and all were accounted for
Comparative	
Notes	Originally, 170 participants were randomised to the PET-CT group. Of those, 21 participants were excluded from the analyses (1 participant refused study investigations after randomisation, and 20 participants did not receive the reference standard) This study was supported by the Ontario Ministry of Health and Long-term Care (grant 06126) and the Canadian Institutes of Health Research (grant MCT-78777) and by Cancer Care Ontario Adverse events: none reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	No		
			Low

Darling 2011 (*Continued*)

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

De Wever 2007

Study characteristics	
Patient sampling	Retrospective consecutive patient series
Patient characteristics and setting	50 participants, median age males = 64 (range = 26 to 83) years/median age females = 60 (range = 46 to 72) years, 44 males/6 females, Belgium **Histology of primary tumour** Adenocarcinoma: N = 23; squamous cell: N = 2; spinocellular carcinoma: N = 14; spinocellular epithelioma: N = 3; carcinoid: N = 1; undifferentiated tumour: N = 1; metastasis: N = 1; benign: N = 5; comorbidities: not reported **Inclusion criteria** Participants who underwent an integrated PET-CT from March 2004 to December 2004 for staging a lung lesion that was suggestive of a lung tumour without metastases on previous clinical or radiological examinations **Exclusion criteria** Participants with evidence of metastatic disease **Previous/all reported tests** Not reported beyond "previous clinical or radiological examinations" **Clinical setting**

	Departments of Radiology/Nuclear Medicine
Index tests	Participants had been instructed to fast for at least 4 hours prior to the FDG PET-CT scan and had blood glucose levels within the normal range. FDG PET-CT scanning was performed using a dual-modality PET-CT tomograph (Biograph LSO Duo; Siemens Medical Solutions). PET imaging was performed 75 min after the administration of 4.5 MBq/kg of FDG. Single-section whole-body spiral CT was performed starting with the head and subsequently covering the neck, thorax, abdomen, and pelvis. 120 ml of a contrast-agent containing 300 mg iodine/ml (Xenetix 300; Guerbet, Sulzbach, Germany) was administered intravenously using an automated injector (1.6 ml/s, scan delay 100 s). CT was performed during breath-hold at expiration tidal volume. A radiologist and nuclear medicine physician interpreted together in consensus the PET-CT fusion images **Covariates** Type of PET-CT scanner: Biograph LSO Duo (Siemens Medical Solutions) FDG dose: 4.5 MBq/kg Injection-to-scan time: 75 min Attenuation correction: yes Cut-off values for test positivity (malignancy): On CT images, lymph node assessment was based on size and consistency of the lymph node. Lymph nodes with a short axis diameter > 10 mm were defined as containing tumour; whereas, enlarged lymph nodes with a lipoid centre were considered benign. On PET images, qualitative analysis of the images was performed by visual identification of areas of increased FDG uptake. A focally increased FDG activity above physiologic levels was considered abnormal and displaying potential malignancy. When there was discordance between PET and CT, the decision about whether the lesion was suspicious or not was made according to the following criteria: 1) a lesion not suggestive on CT but positive on PET was made positive for tumour on integrated PET/CT, 2) pulmonary nodules suggestive for lung metastases on CT but PET-negative were considered as lung metastases, 3) enlarged and suspected lymph nodes on CT but negative on PET were considered as negative on integrated PET-CT, 4) mediastinal hotspots on PET but without a visible lesion on CT were considered as negative (e.g., brown fat tissue) on integrated PET/CT
Target condition and reference standard(s)	Lesion sampling was performed with bronchoscopy and brushing, transbronchial biopsy, transthoracic biopsy, or preoperative biopsy. Surgical staging was performed in all the participants during mediastinoscopy (N = 26) or during surgical exploration (N = 24)
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	No details of funding were reported. However, the study was probably not externally funded because it is retrospective and the data appear to be collected as part of normal practice Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			

De Wever 2007 (*Continued*)

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

El-Hariri 2012

Study characteristics	
Patient sampling	Prospective/retrospective? consecutive? patient series
Patient characteristics and setting	33 participants, median age = 64 (range = 34 to 76) years, 28 males/5 females, Egypt **Histology of primary tumour** Not reported; comorbidities: not reported **Inclusion criteria** Participants with lung cancer who underwent whole body-integrated PET-CT imaging for staging lung lesions in the period from September 2010 till December 2011 **Exclusion criteria** Not reported **Previous/all reported tests** Not reported **Clinical setting** Departments of Radiology Diagnosis and Cardio-thoracic Surgery The inclusion of only participants who received surgery narrows the range of patients who would receive PET/CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	Combined PET-CT imaging was conducted by using the Siemens medical solution system (Siemens Biograph 64 PET-CT scanner). To ensure diagnostic CT image quality, 120 ml of iodinated contrast agent was administered intravenously using an automated injector. CT was performed during breath-hold at expiration tidal volume. This limited breath-hold technique was used to avoid respiration artifacts on the CT images and for a good match between the CT and the PET images. PET imaging was performed 60 min after the administration of 300 MBq (about 8 mci) of FDG by multiple overlapping bed positions (5 min per bed position). Attenuation correction was based on the CT data. Participants had been instructed to fast for 6 h prior to starting the examination. Blood samples collected before the injection of the radioactive tracer ensured blood glucose levels in the normal range **Covariates** Type of PET-CT scanner: Biograph 64 (Siemens Medical Solutions) FDG dose: 300 MBq Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): Lymph nodes with a short axis diameter greater than 10 mm on CT were considered positive. PET images were assessed qualitatively for regions of focally increased FDG uptake, as well as quantitatively by determining standardised uptake values. An increase in FDG uptake to a level greater than that in the surrounding tissue at qualitative analysis and a standard uptake value of more than 2.5 were considered to characterise malignancy. Lymph nodes with increased FDG uptake were considered positive for metastatic spread even when they were smaller than 1 cm in short axis diameter. PET-negative lymph nodes were considered as benign, even when they were larger than 1 cm in short axis diameter. Mediastinal hotspots on PET but without a visible lesion on CT were considered as negative on integrated PET/CT

Target condition and reference standard(s)	Tumour resection and mediastinal lymph node dissection were performed in 22 participants. Surgery was performed within a maximum of 10 days after imaging. The surgeon sampled all visible and palpable lymph nodes that were accessible in the hilum and mediastinum. The remaining 11 participants underwent mediastinoscopy for lymph node staging		
Flow and timing	All the participants were accounted for		
Comparative			
Notes	Funding: no details reported Adverse events: not reported		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	Unclear		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Fischer 2011

Study characteristics	
Patient sampling	Prospective random patient series
Patient characteristics and setting	98 participants, mean age = 62 (range = 42 to 80) years, 53 males/45 females, Denmark **Histology of primary tumour** Adenocarcinoma: N = 29; squamous cell: N = 20; large cell: N = 4; NSCLC other: N = 7; comorbidities: not reported **Inclusion criteria** Participants between 18 and 80 years with newly diagnosed NSCLC that was considered operable after conventional staging procedures **Exclusion criteria** Type 1 diabetes, another malignant condition, confirmed distant metastases, known claustrophobia, and an estimated forced expiratory volume in 1 second of less than 30% after surgery **Previous/all reported tests** CT then FDG-PET/CT followed by invasive diagnostic procedures. Standard staging procedures were governed by local routine based on current guidelines; however, mediastinoscopy was considered mandatory **Clinical setting** Departments of Pulmonology
Index tests	After a 6-hour fast, 400 MBq of (^{18} F)-2-fluoro-deoxy-D-glucose (FDG) was given intravenously, and after a 1-hour rest, the participantwas scanned from the head to the upper thigh with the use of an integrated PET-CT system (GE Discovery LS, GE Healthcare). A diagnostic CT scan, obtained with the use of a standard protocol (80 to 100 mA, 120 kV, a tube-rotation time of 0.5

	second per rotation, a pitch of 6, and a slice thickness of 5 mm, with 70 ml of intravenous contrast medium containing 300 mg of iodine per millilitre (Ultravist, Bayer Schering), administered at a rate of 2.5 ml per second), preceded the PET scan (a 5-minute emission scan per table position and 25 minutes total). The PET scan was reconstructed by filtered back-projection and ordered-subset expectation-maximisation (OS-EM), with data from the CT scan used for attenuation correction. An experienced radiologist and a nuclear medicine specialist evaluated the FDG-PET/CT images side by side, and a consensus was reached on the findings **Covariates** Type of PET-CT scanner: integrated PET-CT scanner (Discovery LS; GE Healthcare) FDG dose: 400 MBq Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): A lesion with increased uptake of FDG in 3 planes when compared with background on a PET scan was classified as malignant. If the image could not be interpreted with confidence, the SUV, defined as the activity per ml within the region of interest divided by the injected dose in megabecquerels per gram of body weight, was calculated, and lesions with a SUV > 2.5 were deemed malignant
Target condition and reference standard(s)	Thoracotomy: N = 60; mediastinoscopy: N = 9; EUS-FNA: N = 19; EBUS-TBNA: N = 4; not applicable: N = 6, as N = 4 had M1 disease, N = 1 had inoperable T4, and N = 1 was N0 or N1, but considered inoperable because of coexisting disease
Flow and timing	All participants were accounted for in the results. 14/98 participants did not receive FDG PET-CT, and 5 participants did not receive the reference standard. Total N reported in the results was therefore = 79
Comparative	
Notes	Funded by the Danish Cancer Society and the Danish Center for Health Technology. The authors stated that the study was subject to no competing interests Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low

Fischer 2011 (*Continued*)

DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Gunluoglu 2011

Study characteristics	
Patient sampling	Retrospective consecutive patient series

Patient characteristics and setting	168 participants, mean age = 60 (range = 30 to 84) years, 149 males/19 females, Turkey **Histology of primary tumour** Adenocarcinoma: N = 57; squamous cell: N = 78; adenosquamous cell carcinoma: N = 9; pleomorphic carcinoma: N = 7; large cell: N = 1; NSCLC not otherwise specified: N = 16; comorbidities: not reported **Inclusion criteria** All NSCLC participants referred to the authors' clinic for surgery between 2007 and 2009 **Exclusion criteria** None listed **Previous/all reported tests** Thoracic CT. PET had been obtained from all participants, and no distant metastasis had been detected **Clinical setting** Departments of Thoracic Surgery It is unclear if the inclusion of only participants who were referred for surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	All the participants were scanned with a multidetector CT-integrated high-resolution PET-CT scanner (Siemens Biograph LSO HI-RES PET/CT). No more information reported **Covariates** Type of PET-CT scanner: multidetector CT-integrated high-resolution PET-CT scanner (Siemens Biograph LSO HI-RES PET/CT) FDG dose: none reported Injection-to-scan time: not reported Attenuation correction: not reported Cut-off values for test positivity (malignancy): When a focus showing elevated FDG uptake in the mediastinum compared with the mediastinal background and adjacent tissues was seen in the PET-CT images, the result was recorded as mediastinal metastasis. There was no prespecified cut-off value
Target condition and reference standard(s)	Cervical mediastinoscopy with (N = 127) or without (N = 41) thoracotomy
Flow and timing	All participants were accounted for in the results. 17/185 participants originally enrolled were excluded (14 in whom PET was obtained using a PET-fusion scanner, 1 refused surgery, 1 high-risk participant could not undergo surgical lung resection, and 1 in whom the interval between PET and mediastinoscopy > 1 month)
Comparative	
Notes	No details of funding were reported. However, the study was probably not externally funded because it is retrospective, and the data appear to be collected as part of normal practice Adverse events: not reported
Methodological quality	

Gunluoglu 2011 (*Continued*)

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Were all patients included in the analysis?	Yes		

Harders 2012

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	114 participants, mean age = not reported (range = not reported) years, gender: not reported, Denmark **Histology of primary tumour** Not reported; comorbidities: not reported **Inclusion criteria** Quote: "Regional patients who were recently diagnosed with NSCLC were prospectively identified for inclusion over a 2- year study period. All patients received a CT as well as an FDG PET/CT examination, and all metastasis suspect lesions were biopsied. Based on all available data, that is the CT, the FDG PET/CT and the biopsy results, a multidisciplinary staging was made: If the patients were staged with T1, N0, M0 disease, they received surgery. If the patients were staged with T2-T4, N0-N3, M0 disease, they received a preoperative mediastinoscopy; if they were eventually staged with T2-T4, N0-N1, M0 disease, they received surgery. In all other instances the patients received oncological treatment" **Exclusion criteria** None listed **Previous/all reported tests** CT **Clinical setting** Departments of Radiology, Nuclear Medicine, Pulmonology, Oncology, Thoracic Surgery, and Pathology
Index tests	FDG PET-CT examinations included the whole body and were performed with an integrated PET-CT scanner (Siemens Biograph 40-slice CT scanner; Siemens Healthcare, Erlangen, Germany). Participants were instructed to submit to 6 hours of fasting prior to the examination. Approximately 400 MBq FDG was injected intravenously. FDG PET-CT scans were performed after a delay of 60 minutes. The FDG PET images were corrected for scatter and iteratively reconstructed. CT acquisition parameters were 40 x 3.0 mm collimation. No contrast medium was administered. 2 consultants in nuclear medicine did the FDG PET-CT reviews. The reviewers were blinded to participant names, participant identifications, and clinical data **Covariates** Type of PET-CT scanner: Siemens Biograph 40-slice CT scanner (Siemens Healthcare, Erlangen, Germany) FDG dose: circa 400 MBq Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): FDG uptake was compared with the background uptake of the liver. Thus, lymph node uptake was rated on a scale of 1 to 3: (1) no uptake, (2)

	probably increased uptake (i.e., below liver level uptake), (3) definitely increased uptake (i.e., at or above liver level uptake). A rating of 1 was considered normal; a rating of 2 or 3 was considered abnormal
Target condition and reference standard(s)	Tissue sampling from the participants' mediastinums. In participants who did not receive surgery, tissue sampling was obtained by preoperative mediastinoscopy with sampling from nodal stations 1, 2R/L, 3A, 4R/L, and 7; if necessary, it was obtained by anterior mediastinotomy from nodal stations 5 and 6. All mediastinoscopies/-tomies (N = 25) were guided by both CT and FDG-PET/CT examinations. In participants who received surgery (N = 89), tissue sampling was obtained by complete lymph node resection (i.e., resection of all visible and palpable mediastinal and hilar lymph nodes from nodal stations 2R; 4R; 7, 8, 9, 10, 11+ for right-sided tumours; and 5, 6, 7, 8, 9, 10, 11+ for left-sided tumours)
Flow and timing	All participants received the reference standard and were included in the analyses
Comparative	
Notes	Funding: not reported Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Unclear		
Was a positive result defined?	Yes		

			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Hu 2011

Study characteristics	
Patient sampling	Patient series
Patient characteristics and setting	102 participants, mean age = 56 (SD = 14) years, 79 males/23 females, China **Histology of primary tumour** Adenocarcinoma: N = 41; squamous cell: N = 52; NSCLC not otherwise specified: N = 9; comorbidities: pulmonary (diagnosed on the basis of chest CT or PET/CT images): N = 53, including obstructive pneumonia (N = 24), unspecified acute or chronic infection pneumonia (N = 16), interstitial pneumonitis (N = 3), previous pulmonary tuberculosis (N = 4), active pulmonary tuberculosis (N = 1), atelectasis (N = 2), aspergillosis (N = 1), pneumoconiosis (N = 1), and bronchiectasis (N = 3) **Inclusion criteria** From March 2004 to March 2009, participants with pathologically proven NSCLC who received dual-time-point FDG PET-CT scanning before radical surgery were included in this study. It is unclear if enrolment was prospective or retrospective or of consecutive participants **Exclusion criteria**

	None listed **Previous/all reported tests** All participants underwent conventional lung cancer staging on the basis of clinical information and FDG PET-CT studies **Clinical setting** Departments of Thoracic Surgery, Radiation Oncology, Oncology, and Nuclear Medicine The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	All participants were fasted for at least 6 hours before the examination. PET images were acquired using an integrated PET-CT scanner (Discovery LS; GE Healthcare, Milwaukee, WI, USA). The dosage of the FDG injection was 370 MBq. Immediately after unenhanced CT, a PET emission scan was performed that covered the identical transverse field of view. Images were acquired twice: an early scan that included the head, thorax, abdomen, pelvis, and thigh approximately 60 minutes after FDG injection (range = 50 to 65 minutes); and a delayed scan that included the chest approximately 120 minutes after injection (range = 110 to 140 minutes). The acquisition parameters for these 2 scans were the same. 2 experienced nuclear medicine physicians who were unaware of the participant's clinical history and the results of previous conventional imaging tests reviewed the PET-CT images. On a transaxial slice of attenuation corrected PET, standardised uptake value (SUV) was obtained by placing regions of interest on the primary tumours and the LNs in each station that had been identified by visual analysis. To minimise partial volume effects, the maximum SUV (SUVmax) within an region of interest that had been automatically calculated by the Xeleris software was used. The SUVmax of the early and delayed image were defined as SUVe and SUVd, respectively **Covariates** Type of PET-CT scanner: Discovery LS (GE Healthcare, Milwaukee, WI, USA) FDG dose: 370 MBq Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): The retention index (RI) was the percentage change of SUV in the respective lesions between early and delay images: RI = (SUVd - SUVe)/SUVe X 100%. SUVmax $\geq$ 2.5 were considered to be positive metastatic LNs on single-time-point imaging; a RI of 10% is the criterion for metastatic LNs on dual-time-point imaging
Target condition and reference standard(s)	All participants received radical surgery with system mediastinal LN dissection within 4 days (range = 3 to 7 days) following the PET-CT scanning. A pathologist examined all lymph nodes for the presence/absence of malignancy
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	Suported by the National High Technology Research and Development Program of China Adverse events: not reported **Single-time-point data** Participants with a comorbidity: TP = 14, FP = 15, FN = 2, TN = 22 Participants without a comorbidity: TP = 15, FP = 10, FN = 3, TN = 21 **Dual time-point data**

	Participants with a comorbidity: TP = 15, FP = 12, FN = 1, TN = 25 Participants without a comorbidity: TP = 15, FP = 9, FN = 3, TN = 22		

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			

Hu 2011 (*Continued*)

Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Hwangbo 2009

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	117 participants, median age = 66 (range = 41 to 84) years, 92 males/25 females, South Korea **Histology of primary tumour** Adenocarcinoma: N = 55; squamous cell: N = 53; large cell: N = 7; sarcomatoid carcinoma: N = 1; NSCLC NOS: N = 1; comorbidities: not reported **Inclusion criteria** Participants with histologically confirmed or strongly suspected potentially operable NSCLC from October 2006 to October 2007. Participants were required to have at least 1 mediastinal lymph node in an accessible location by EBUS-TBNA, with a short diameter of 5 to 20 mm on chest CT scan axial image **Exclusion criteria** Participants with Pancoast tumours or unresectable tumours detected by white light bronchoscopy, medical inoperable participants, and participants not considered physically fit for surgery. Participants who had M1 disease, inoperable T4 disease, a bulky mediastinal lymph node (short diameter of 2 cm on chest CT scan axial image), or extra-nodal invasion of the mediastinal lymph node visible on chest CT scan. When an abnormal supraclavicular lymph node was detected by chest CT scan or integrated PET-CT scan in otherwise eligible participants, fine-needle aspiration was performed, and participants who were found to have supraclavicular lymph node metastasis were excluded **Previous/all tests** Surgical tumour resectability was evaluated after staging workup for NSCLC, including either a CT scan of the chest and upper abdomen, integrated PET/CT scan, brain MRI, bone scan, or a combination of the aforementioned **Clinical setting** Center for Lung Cancer
Index tests	Participants fasted for 8 h and then received an I injection of FDG (10 to 15 mCi). Scanning was performed 60 min later PET/CT scan images were obtained by using either 1 of the following 2 combined PET-CT scanners: a Biograph LSO (Siemens Medical Solutions; Hoffman Estates, IL, USA) or a Discovery LS (GE Medical Systems; Milwaukee, WI, USA). On each PET/CT scan, a spiral CT scan was performed and integrated with PET scan images. The data were analysed using dedicated workstations loaded

	with e.soft (Siemens Medical Solutions) and eNTEGRA™ (GE Medical Systems) software. The SUV was calculated as follows: SUV (decay-corrected activity (in kilobecquerels) per ml of tissue volume)/injected-FDG activity (in kilobecquerels)/body mass (in grams). The SUV was obtained by locating a region of interest on a lesion, and the maximum SUV within a region of interest was used. A maximum SUV > 2.5 on a lymph node was interpreted as positive **Covariates** Type of PET-CT scanner: a Biograph LSO (Siemens Medical Solutions, Hoffman Estates, IL, USA) or a Discovery LS (GE Medical Systems; Milwaukee, WI, USA) FDG dose: 10 to 15 mCi Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): maximum SUV > 2.5
Target condition and reference standard(s)	The reference standard consisted of EBUS-TBNA for all 117 participants and surgical lymph node dissection in the 92 participants who were believed to be N2-negative on the basis of EBUS-TBNA. The results reported were based on the composite reference standard of both EBUS-TBNA and surgical lymph node dissection
Flow and timing	12 of the originally 129 enrolled participants were excluded from the analyses: 2/12 participants had SCLC, 1/12 participants had organising pneumonia, 7/12 participants refused surgery, and 2/12 participants did not undergo lymph node dissection because of unexpected pleural metastasis found during surgery
Comparative	
Notes	This work was supported by a National Cancer Center grant. The authors reported to the American College of Chest Physicians that no significant conflicts of interest exist with any companies/organisations whose products or services may be discussed in their article Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			

Hwangbo 2009 (*Continued*)

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Iskender 2012

Study characteristics	
Patient sampling	Prospective? consecutive patient series
Patient characteristics and setting	286 participants, mean age = 58.5 (SD = 9.3, range = 33 to 81) years, 362 males/24 females, Turkey **Histology of primary tumour** Adenocarcinoma: N = 90; squamous cell carcinoma: N = 158; large cell: N = 5; adenosquamous carcinoma: N = 4; carcinosarcoma: N = 3; spindle cell carcinoma: N = 1; NSCLC NOS: N = 25;

	comorbidities: not reported **Inclusion criteria** From September 2005 to March 2009, consecutive participants with NSCLC histology were imaged with PET-CT within 90 days before mediastinoscopy, thoracotomy, or both **Exclusion criteria** Neoadjuvant chemotherapy (N = 22), previous history of NSCLC (N = 9), or other malignancies within 5 years (N = 11) and clinical stage IV **Previous/all tests** None reported **Clinical setting** Department of Thoracic Surgery
Index tests	PET/CT images were obtained at 10 different centres, all of which used multidetector CT-integrated PET scanners. 225 participants were imaged at 4 different imaging centres that used the same detector (Siemens Biograph LSO HI-RES PET/CT; Siemens AG, Erlangen, Germany). All integrated PET-CTs were performed with participants fasting for at least 6 hours, and the blood glucose level was below 8.3 mmol/l before FDG injection. Whole-body scans were obtained 60 min after intravenous injection of 10 to 20 mCi FDG. For PET-CT imaging, simultaneously acquired CT data were used to correct attenuation. Scans from centres other than those using the Siemens PET-CT scanner were also eligible. Nuclear medicine physicians experienced in interpreting PET-scans visually evaluated images acquired to detect mediastinal metastasis. In the visual evaluation, FDG uptake was considered to be positive in the evaluation of mediastinal lymph nodes if tracer activity was significantly higher than mediastinal background activity. The SUVmax for all primary tumours and positive lymph nodes was provided. A PET-CT scan was interpreted as positive if the SUVmax of mediastinal lymph nodes exceeded 2.5 **Covariates** Type of PET-CT scanner: Biograph LSO HI-RES (Siemens AG, Erlangen, Germany) FDG dose: 10 to 20 mCi Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): maximum SUV > 2.5
Target condition and reference standard(s)	Mediastinal lymph node staging was completed in all participants by means of standard cervical mediastinoscopy, extended cervical mediastinoscopy, thoracotomy, or a combination of the aforementioned. Participants with negative mediastinoscopy underwent resection and systematic lymph node sampling. Participants with positive mediastinoscopy were referred to the oncology clinic for neoadjuvant or definitive treatment
Flow and timing	All participants were accounted for. The mean time interval between PET-CT and surgical staging was 16.3 (SD = 11.5, range = 2 to 90) days
Comparative	
Notes	No details of funding were reported. The authors declared no conflicts of interest Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Iskender 2012 (*Continued*)

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Were all patients included in the analysis?	Yes		

Jeon 2010

Study characteristics	
Patient sampling	Retrospective case-control study; however, the data were collapsed across categories as both cases and controls had NSCLC
Patient characteristics and setting	Cases: 42 participants, mean age = 66 (SD = 5.2) years, 40 males/2 females, South Korea **Histology of primary tumour** Adenocarcinoma: N = 23; squamous cell: N = 12; NSCLC other: N = 7; comorbidities: idiopathic pulmonary fibrosis: N = 42; none had sarcoidosis Controls: 168 participants, mean age = 65 (SD = 5.6) years, 130 males/38 females, Korea **Histology of primary tumour** Adenocarcinoma: N = 92; squamous cell: N = 48; NSCLC other: N = 28 **Inclusion criteria** Participants who had undergone surgical nodal staging or curative resection for NSCLC after preoperative thoracic CT and whole-body PET/CT examinations from March 2003 to December 2008 Cases: participants with NSCLC and idiopathic pulmonary fibrosis (IPF) Controls: participants with NSCLC, but not IPF who were age- and histopathologic type-matched to the cases **Exclusion criteria** Participants who received neoadjuvant therapy of any kind Cases: participants whose pulmonary diagnosis could be related to asbestos and other environmental exposure or to the presence of underlying collagen vascular disease and participants who had received corticosteroid treatment within 2 months of surgery **Previous/all tests** None listed other than contrast-enhanced CT and integrated PET-CT scan **Clinical setting** Secondary/tertiary setting
Index tests	Peripheral blood glucose $\leq$ 150 mg/dL in all participants. Participants received an IV injection of 370 MBq (10 mCi) of FDG, which was followed by 45 to 60 minutes of rest before scanning. Image acquisition was acquired using a Discovery LS PET-CT device (GE Healthcare). CT was performed according to a standard protocol with the following parameters: 140 kV; 80 mAs; tube rotation time, 0.5 second per rotation; pitch, 6; and section thickness, 5.0 mm (to match the PET section thickness). Immediately after an unenhanced CT scan was obtained, PET was performed in an identical axial field of view. PET-image data sets were reconstructed iteratively using the ordered subsets expectation maximisation algorithm and by application of segmented attenuation correction to the CT data. Integrated PET-CT images were evaluated jointly and prospectively by a chest radiologist (with 20 years of CT interpretation experience and 6 years of PET-CT analysis) and a nuclear medicine physician (with 15 years of experience of general nuclear medicine interpretation and 6 years of PET-CT interpretation experience). Both were unaware of clinical or histopathologic

	results **Covariates** Type of PET-CT scanner: integrated PET-CT scanner (Discovery LS; GE Healthcare) FDG dose: 370 MBq (10 mCi) Injection-to-scan time: 45 to 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): When lymph nodes showed increased FDG uptake (> mediastinal blood pool uptake) in the thorax, they were considered as metastatic irrespective of their size. However, even though lymph nodes showed higher FDG uptake than the mediastinal blood pool uptake on PET component images of PET-CT, when they showed higher attenuation (≥ 70 HU using a region of interest-based measurement) than the surrounding great vessels or benign calcifications (central, nodular, diffuse, laminated, or popcornlike) on CT component images of PET-CT, the lymph nodes were designated as benign. Enlarged lymph nodes with a short axis diameter > 10 mm were regarded as benign when the nodes had less FDG uptake than that of mediastinal blood pool at PET-CT. There was no prespecified cut-off value
Target condition and reference standard(s)	Cases: Surgical staging included thoracotomy alone (N = 21) or both mediastinoscopy and thoracotomy (N = 21). The surgical procedures performed were lobectomy (N = 38), bilobectomy (N = 1), and segmentectomy or wedge resection (N = 3) Controls: Mediastinal lymph nodes were obtained by using thoracotomy alone (N = 46), mediastinoscopy plus thoracotomy (N = 111), or mediastinoscopy alone (N = 11). Surgical resection included lobectomy (N = 134), bilobectomy (N = 13), segmentectomy or wedge resection (N = 1) , and pneumonectomy (N = 9). For 11 participants, only mediastinoscopic nodal staging results were available as curative resection was deferred because of the pathologically proven N2 disease, indicating the need for neoadjuvant concurrent chemoradiation therapy (n = 8), or because participants declined surgery (n = 3)
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	No details of funding were reported. However, the study was probably not externally funded because it is retrospective, and the data appear to be collected as part of normal practice. Please note, 42/210 participants also had idiopathic pulmonary fibrosis Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		

Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Kim 2007

Study characteristics	
Patient sampling	Prospective consecutive? patient series
Patient characteristics and setting	674 participants, mean age = 61 (range = 30 to 90) years, 502 males/172 females, South Korea **Histology of primary tumour** Adenocarcinoma: N = 333; squamous cell: N = 271; bronchioloalveolar carcinoma: N = 14; adenosquamous: N = 7; large cell neuroendocrine carcinoma: N = 29; sarcomatoid carcinoma: N = 15 (pleomorphic carcinoma: N = 14, spindle cell carcinoma: N = 1); NSCLC NOS: N = 5; comorbidities: not reported, but 218 (of the 674) participants had a past medical history of pulmonary tuberculosis (as determined at clinical or imaging studies) **Inclusion criteria** Participants referred for surgery for NSCLC between March 2003 and March 2006 **Exclusion criteria** None listed, but participants were excluded because conventional staging studies or integrated whole-body PET-CT suggested extra-thoracic metastasis (N = 41), because they received chemotherapy (N = 6) or chemoradiotherapy (N = 20) before surgical staging at another hospital, or because they had carcinoids (N = 5) or salivary gland type tumours (N = 10) **Previous/all tests** Conventional lung cancer staging on the basis of clinical information, stand-alone chest CT with intravenous injection of 100 mL of iopamidol (Iopamiron 300; Bracco, Milan, Italy), and an integrated whole-body PET-CT study **Clinical setting** Tertiary setting The inclusion of only participants who were referred for surgery and the exclusion of participants with extra-thoracic metastasis based on PET-CT narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	All participants fasted for at least 6 hours before the PET-CT examination, although oral hydration with glucose-free water was allowed. After a normal blood glucose level in the peripheral blood was ensured, participants received an intravenous injection of 370 MBq (10 mCi) of FDG and then rested for 45 min before undergoing imaging. Image acquisition was performed with an integrated PET/CT device (Discovery LS, GE Medical Systems). The participants were scanned from the head to the pelvic floor and were allowed normal shallow respiration during the acquisition of the non-contrast-enhanced CT scans. PET-image data sets were reconstructed iteratively using the ordered subsets expectation maximisation algorithm and by applying the segmented measured attenuation correction to the CT data. Coregistered images were displayed by using software (eNTEGRA; GE Medical Systems) that allowed image fusion and analysis. 1 chest radiologist and 1 nuclear medicine physician, both unaware of clinical, stand-alone CT and pathologic results, together prospectively evaluated integrated PET-CT datasets. Decisions were reached by consensus **Covariates** Type of PET-CT scanner: integrated PET-CT device (Discovery LS; GE Medical Systems) FDG dose: 370 MBq (10 mCi) Injection-to-scan time: 45 min Attenuation correction: yes, segmented measured attenuation correction to the CT data Cut-off values for test positivity (malignancy): Mediastinal nodes with an increased glucose uptake and a distinct margin were considered positive for malignancy. Nodes were regarded as having an increased glucose uptake when they demonstrated FDG uptake at a level greater than that of the

	surrounding mediastinal tissue. Calcification was regarded as present when nodular, laminated, or diffuse and when the attenuation was 200 household unit (HU). A highly attenuating node was defined as 1 that appeared to have a higher attenuation than mediastinal vascular structures with an attenuation of 70 HU using a region of interest (ROI)-based measurement. Even if glucose uptake was high (higher than the background activity), calcified lymph nodes or lymph nodes with a higher attenuation than surrounding great vessels on the CT images of integrated PET-CT were regarded as benign
Target condition and reference standard(s)	Surgical staging included mediastinoscopy (N = 121), mediastinoscopy plus thoracotomy (N = 309) , and thoracotomy alone (N = 244). In 121 participants, only mediastinoscopic nodal staging results were available because curative resection was deferred on account of the presence of positive nodes indicating neoadjuvant concurrent chemoradiation therapy (N = 108) or because participants denied surgery (N = 13) in spite of negative nodes on mediastinoscopic evaluation. Tumour resection and extensive mediastinal lymph node dissection with thoracotomy were performed on 553 participants. During mediastinoscopy, the 2R, 4R, 2L, 4L, and 7 ATS lymph node map areas were routinely sampled, and during thoracotomy, according to the routine surgical protocol, surgeons dissected all visible and palpable lymph nodes accessible in the mediastinum irrespective of size. Specifically, all encountered lymph nodes were removed from 10R, 9, 8, 7, 4R, 3, and 2R, the ATS lymph node map areas for tumours of the right lung, and from areas 10L, 9, 8, 7, 6, 5, and 4L for the left lung. When necessary, especially when an imaging study suggested possible nodal metastasis in other nodal stations other than those included in routine lymph node dissection, group 1 (highest mediastinal) or 2L (when tumours were located in the left lung) nodes were also evaluated during mediastinoscopy or thoracotomy. In the 553 participants of the surgical tumour resection group, 389 participants underwent lobectomy, 72 participants underwent bilobectomy, 14 participants underwent sleeve lobectomy, and 78 participants underwent pneumonectomy. The mean time interval between the index test and reference standard was 13 days (range = 1 to 42 days; median = 7 days)
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	No details of funding were reported. However, the study was probably not externally funded because although this data collection was prospective, it appears to be collected as part of normal practice This study was performed in a tertiary referral centre of a TB-endemic country, where TB is still a serious public health problem and the incidence of active TB was as high as 73 per 100,000 population (intermediate burden country according to World Health Organization classification) in 2005 At least, 38 of the 674 participants belonged to the patient population of Shim 2005 (as reported in Kim 2006) Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Koksal 2013

Study characteristics	
Patient sampling	Retrospective patient series
Patient characteristics and setting	81 participants, mean age = 59.8 (SD = 8, range = 38 to 74) years, 77 males/4 females, Turkey **Histology of primary tumour** Adenocarcinoma: N = 32; squamous cell: N = 43; adenosquamous: N = 4; pleomorphic carcinoma: N = 2; comorbidities: not reported **Inclusion criteria** Patients with NSCLC who had not received chemotherapy or radiotherapy, with a PET-CT examination at the time of initial staging who subsequently underwent surgical resection **Exclusion criteria** None listed **Previous/all tests** Not reported **Clinical setting** Tertiary setting The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	PET-CT was carried out with an integrated PET/CT scanner (Siemens, Biograph-6 True Point) within the 30 days before the surgery in all of the participants. Participants were instructed to fast for at least 6 hours before the examination. After confirmation of a normal peripheral blood glucose level (< 180 mg/dL), the participants received an intravenous injection of 145 μCi/kg (maximum 200 μCi) of FDG and rested for 60 minutes before the scan. Images were obtained from the base of skull to mid-thigh level. Software determined automatically the SUVmax of the primary tumours and each suspicious lymph node stations after delineation of the region of interest on attenuation-corrected PET-CT images. All PET-CT scans were re-evaluated. SUVmax of the primary tumours and dissected mediastinal and hilar lymph node stations were noted. Positivity of lymph node stations was rated according to 2 criteria: 1) SUVmax > 2.5; 2) FDG uptake higher than the surrounding mediastinal blood pool **Covariates** Type of PET-CT scanner: Siemens, Biograph-6 True Point FDG dose: 145 μCi/kg (maximum 200 μCi) (equivalent to 5.365 MBq/kg up to a max of 740 MBq) Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): SUVmax > 2.5, but also data for lymph node stations were considered as positive if there was a FDG uptake higher than the surrounding mediastinal blood pool
Target condition and reference standard(s)	Resection (lobectomy: N = 55, pneumonectomy: N = 25, and wedge resection: N = 1) with complete ipsilateral hilar and mediastinal lymph node dissection

Flow and timing	All participants were accounted for in the data. Index test received within 30 days of reference standard
Comparative	
Notes	The author emailed the individual participant data, which we classified according to 2 criteria: 1) SUVmax > 2.5: TP = 8, FN = 4, FP = 22, TN = 47; 2) FDG uptake higher than the surrounding mediastinal blood pool: TP = 8, FN = 4, FP = 31, TN = 38. The former data have been used for analysis Funding: no details reported, but the authors did state they had no financial conflict of interest that could bias the work Adverse events: none reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Kuo 2012

Study characteristics	
Patient sampling	Retrospective consecutive patient series
Patient characteristics and setting	102 participants, mean (?) age = 63.1 (range = 34 to 81) years, 56 males/46 females, Taiwan **Histology of primary tumour** Adenocarcinoma: N = 74; squamous cell carcinoma N = 14; other N = 14; comorbidities: not reported **Inclusion criteria** All newly diagnosed NSCLC participants who underwent PET-CT < 1 month before surgery between May 2006 and October 2010 **Exclusion criteria** Participants who underwent mediastinoscopy alone (N = 20) or lymph node sampling during operation (N = 34), and participants who had been treated with chemotherapy (N = 12) or targeted therapy (N = 6) before PET-CT **Previous tests** None listed **Clinical setting** Thoracic Surgery Service The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	**Covariates** Type of PET-CT scanner: integrated PET/CT scanner (Discovery ST16; GE Medical Systems,

	Milwaukee, WI, USA) FDG dose: 370 to 555 MBq Injection-to-scan time: 50 min Attenuation correction: yes Cut-off values for test positivity (malignancy): 2 experienced nuclear medicine physicians interpreted the images by consensus. Lymph nodes were considered positive for metastasis when FDG uptake was higher than the mediastinal blood pool. Lymph nodes with FDG uptake higher than the mediastinal blood pool but distributed symmetrically in the bilateral mediastinal and hilar areas were considered negative for metastasis. For each participant, the lymph nodes considered positive for metastasis by visual interpretation were selected for SUV measurement. In addition, lymph node stations that had pathologically confirmed metastatic foci, but without abnormal FDG uptake were also measured. In participants who were considered negative by visual interpretation and without pathological evidence of N2 disease, SUV was measured in regions encompassing the first ipsilateral mediastinal lymph node area that lay on the drainage area of the primary tumour lobe. The mean SUV of the mediastinal blood pool and the liver were obtained from the aortic arch and the hepatic parenchyma, respectively. Node SUVs were then divided by the mean SUV of the aortic arch and liver to calculate the node/aorta and node/liver SUV ratios
Target condition and reference standard(s)	Histology from surgical resection of the primary tumour and standard mediastinal lymph node dissection
Flow and timing	All participants were accounted for in the results. All participants underwent surgical staging. There were no withdrawals
Comparative	
Notes	This study was supported by grant NSC-97-2314-B-182A-101-MY3 from the National Science Council of Taiwan Adverse events: none reported Node SUV cut-off 3.15: TP = 13, FN = 6, FP = 20, TN = 63 Node/aorta SUV ratio cut-off 1.37: TP = 17, FN = 2, FP = 33, TN = 50 Node/liver SUV ratio cut-off 1.02: TP = 16, FN = 3, FP = 28, TN = 55 It was the visual interpretation results currently included under test data

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		

			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Lee 2007

Study characteristics	
Patient sampling	Retrospective patient series
Patient characteristics and setting	126 participants, mean age = 67 (range = 37 to 86) years, 53 males/73 females, USA **Histology of primary tumour** Not reported; comorbidities: not reported **Inclusion criteria** All newly diagnosed participants with biopsy-proven NSCLC who underwent surgical mediastinal lymph node biopsy by cervical mediastinoscopy, anterior mediastinotomy, thoracotomy, or a combination of these methods, between January 1995 and December 2005 on the Thoracic Surgery Service at the University of California, Davis Cancer Center, after a preoperative staging integrated FDG PET-CT scan **Exclusion criteria** None listed **Previous tests** None listed **Clinical setting** Thoracic Surgery Service It is unclear if the inclusion criteria narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease on PET-CT
Index tests	PET-CT images were obtained after participant fasting for a minimum of 4 hours with an integrated PET/CT scanner (Discovery LS; GE Medical Systems, Waukesha, WI, USA; or ECAT Reveal XVI; CTI, Knoxville, TN, USA). Whole-body scans were obtained 30 to 60 minutes after intravenous injection of 10 to 20 mCi of FDG. Simultaneously acquired CT data were used for attenuation correction. All studies were read by dedicated nuclear medicine physicians with a specialty in interpreting PET scan images. Clinical histories and pertinent CT scans were available for review **Covariates** Type of PET-CT scanner: PET-CT images were obtained with an integrated PET-CT scanner (Discovery LS; GE Medical Systems, Waukesha, WI, USA; or ECAT Reveal XVI; CTI, Knoxville, TN, USA) FDG dose: 10 to 20 mCi Injection-to-scan time: 30 to 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): Mediastinal lymph nodes were read as positive if their activity was definitely above the surrounding mediastinal activity and not according to standard uptake values (SUV). There was no prespecified cut-off value
Target condition and reference standard(s)	Extended mediastinal lymph node staging was completed in all participants by cervical mediastinoscopy, anterior mediastinotomy, or thoracotomy. In participants with normal mediastinoscopy results, thoracotomy followed typically within 14 days. All visible and technically feasible lymph nodes were removed. Only participants with pathologic disease in lymph nodes that would have been accessible by mediastinoscopy (stations 2, 4, and 7), mediastinotomy (stations 5 and 6), right thoracotomy (stations 2, 4, 7, and 9), or left thoracotomy (stations 4, 5, 6, and 9) were considered positive in this study. The results of PET and CT scanning were available to the surgeon at the time of resection. The average length of time between PET-CT and resection was 33.8 days

Flow and timing	All participants were accounted for in the results. All participants underwent surgical staging. There were no withdrawals
Comparative	
Notes	No details of funding were reported. However, the study was probably not externally funded because the data are retrospective and apparently collected as part of normal practice Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge	Unclear		

of the results of the index tests?			
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Lee 2009a

Study characteristics	
Patient sampling	Prospective consecutive? patient series
Patient characteristics and setting	182 participants, mean age = 60.7 (SD = 10.8) years, 126 males/56 females, South Korea **Histology of primary tumour** Adenocarcinoma: N = 93; squamous cell: N = 66; bronchioloalveolar carcinoma: N = 5; large cell carcinoma: N = 7; cancer NOS: N = 11; comorbidities: not reported **Inclusion criteria** Participants who underwent preoperative FDG PET-CT and subsequent surgical resection of NSCLC between March 2004 and February 2006 **Exclusion criteria** None listed, but participants with metastatic lesions on preoperative PET-CT images and participants who had had neoadjuvant chemotherapy or radiotherapy for contralateral or bulky mediastinal node metastases before thoracotomy or mediastinoscopy were excluded **Previous/all tests** All participants underwent a contrast-enhanced CT scan of the thorax and FDG PET-CT scan as part of staging work-up. No further information provided **Clinical setting** Secondary setting The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	PET-CT was performed using a Gemini PET-CT system (Philips, Milpitas). All participants fasted for at least 6 h before the PET-CT scan, and only glucose-free water was allowed. An intravenous injection of 5.18 MBq of FDG/kg of body weight was administered, and participants rested for 60 min before imaging. PET-CT data were obtained with participants in the supine position. Attenuation correction was done based on CT data. 2 experienced nuclear medicine physicians

	evaluated all PET-CT images. During evaluation of CT images, the short axis of mediastinal lymph nodes was measured and positive nodes were defined as those with a short axis diameter greater than 1 cm. In addition to size, the presence of calcification was considered on non-contrast CT images **Covariates** Type of PET-CT scanner: Gemini PET-CT system (Philips, Milpitas) FDG dose: 5.18 MBq of FDG/kg Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): Integrated PET-CT images were evaluated visually. The maxSUVs were measured in all lymph nodes with increased FDG uptake. First, mediastinal lymph nodes with focally increased FDG uptake higher than mediastinal blood pool uptake were judged as positive, taking the SUVs of the lymph nodes into consideration. On a second interpretation of PET-CT images, calcified high-attenuation lymph nodes (defined as nodes with higher attenuation than that of the mediastinal vascular structures and more than 70 HU on non-contrast CT images) were interpreted as benign, irrespective of FDG uptake. Furthermore, bilateral symmetric paratracheal nodes on FDG PET images with a hilar or interlobar nodal distribution with similar FDG uptake or mediastinal nodes with a symmetric hilar or interlobar nodal distribution with similar FDG uptake were also judged as benign. Even though the attenuation of some lymph nodes with a typical distribution pattern was lower than 70 HU, all the lymph nodes in the participants with this pattern were interpreted as benign
Target condition and reference standard(s)	Thoracotomy was performed in 169 of 182 participants. Mediastinoscopic biopsy without thoracotomy was performed in the remaining 13 participants because of a pathological high N-stage found on mediastinoscopic biopsy. An additional mediastinoscopic biopsy was performed in 31 of the 169 participants in whom thoracotomy was performed, and the remaining 138 participants had thoracotomy only. All the mediastinal nodes that were positive on FDG PET-CT images or on contrast-enhanced CT images were sampled or dissected at thoracotomy, mediastinoscopic biopsy, or both
Flow and timing	All participants were accounted for in the results. All participants underwent surgical staging. There were no results that uninterpretable results
Comparative	
Notes	No details of funding were reported. However, the study was probably not externally funded because although this data collection was prospective, it appears to be collected as part of normal practice The participant population comes from a region with a high prevalence of granulomatous disease Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Lee 2011

Study characteristics	
Patient sampling	Retrospective patient series
Patient characteristics and setting	54 participants, median age = 66 (range = 45 to 83) years, 48 males/6 females, South Korea **Histology of primary tumour** Not reported. Comorbidities: diabetes: N = 7; hypertension: N = 15; chronic renal failure: N = 1; gastric ulcer or early gastric cancer: N = 4; chronic obstructive pulmonary disease: N = 2; history of tuberculosis: N = 14; stage IA to IIB: N = 36; stage IIIA: N = 10; stage IIIB: N = 4; other cancer metastasis to the lung: N = 4 **Inclusion criteria** Participants who had undergone both chest CT and FDG PET-CT before surgical resection with at least ipsilateral 4- and 7-lymph node dissection for the treatment of primary or metastatic lung cancer between January 2004 and December 2006 at the Seoul National University Hospital, and who had radiographic TB sequelae ipsilateral to the resected lung in the form of fibrotic bands, small calcified nodules, or bronchiectasis in the upper lobes observed on chest CT preoperatively **Exclusion criteria** None listed **Previous/all tests** None listed apart form CT **Clinical setting** Tertiary referral hospital
Index tests	On CT, mediastinal lymph node enlargement (i.e., positive) was defined as the presence of lymph nodes larger than 1 cm in their smallest diameter **Covariates** Type of PET-CT scanner: not reported FDG dose: not reported Injection-to-scan time: not reported Attenuation correction: not reported Cut-off values for test positivity (malignancy): On FDG PET-CT, mediastinal nodes with increased glucose uptake satisfying both qualitative (greater than that of the surrounding tissue) and quantitative (a maximum standardised uptake value (SUV) adjusted for participant body weight of $\geq$ 3.0 with a distinct margin) criteria were considered positive
Target condition and reference standard(s)	Pathologic findings in resected specimens
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	No details of funding were reported. However, the study was probably not externally funded because the data are retrospective and apparently collected as part of normal practice Population of TB participants Adverse events: not reported
Methodological quality	

Lee 2011 (*Continued*)

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Lee 2011 (*Continued*)

Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Lee 2012

Study characteristics	
Patient sampling	Retrospective consecutive? patient series
Patient characteristics and setting	160 participants, mean age = 60 (range = 29 to 80) years, 62 males/98 females, South Korea **Histology of primary tumour** Adenocarcinoma with bronchioloalveolar cell carcinoma: N = 55; adenocarcinoma with mixed type: N = 80; bronchioloalveolar cell carcinoma: N = 25; comorbidities: not reported **Inclusion criteria** Participants with pathologically confirmed T1 NSCLC between January 2005 and May 2011 who had available FDG PET-CT and thin-section chest CT (slice thickness ≤ 2.5 mm) before treatment, an interval ≤ 2 months between FDG PET-CT, CT and treatment, NSCLC appearing as subsolid nodules on CT with lymph node staging, no previous chemotherapy/radiotherapy, and no previous/concurrent malignancy **Exclusion criteria** None listed **Previous/all tests** FDG-PET/CT and CT. No further details reported **Clinical setting** Secondary/tertiary setting
Index tests	Before intravenous administration of FDG (5.2 MBq/kg body weight), all participants fasted for ≥ 6 hours. After administration, participants rested for 60 min before imaging. Thereafter, whole-body PET images were acquired with the conventional protocol of FDG PET using a Gemini (Philips Medical Systems, Cleveland, OH, USA) equipped with a 2-slice CT or Biograph 40 (Siemens Medical Solutions, Knoxville, TN, USA). 2 nuclear medicine physicians with 9 and 3 years PET/CT experience, respectively, evaluated all FDG PET-CT images, and all decisions were reached in consensus. SUVmax of the primary lesions was also calculated. SUVmax threshold cut-off of 3.5 was determined according to the previous experience of the authors' institute and other reports in a tuberculosis-endemic area **Covariates** Type of PET-CT scanner: Gemini (Philips Medical Systems, Cleveland, OH, USA) equipped with a 2-slice CT or Biograph 40 (Siemens Medical Solutions, Knoxville, TN, USA) FDG dose: 5.2 MBq/kg Injection-to-scan time: 60 min Attenuation correction: not reported Cut-off values for test positivity (malignancy): All lymph nodes in the thorax and extra-thoracic regions with abnormal FDG uptake (SUVmax > 3.5) were considered positive, unless they showed

	high attenuation (> 70 HU) or benign calcification (central nodular, laminated, popcorn, or diffuse) on unenhanced CT images
Target condition and reference standard(s)	Pathological results from thoracotomy with (N = 128) or without (N = 32) mediastinoscopy. During thoracotomy, all visible and palpable lymph nodes accessible in the mediastinum were dissected. When preoperative imaging results suggested possible lymph node metastasis, they were also evaluated during mediastinoscopy or thoracotomy. Contralateral hilar lymph node metastasis was determined using clinical and imaging follow-up studies
Flow and timing	All participants were accounted for in the results. All participants underwent surgical staging. There were no uninterpretable results
Comparative	
Notes	From a tuberculosis-endemic area The Research Grant of Korea Foundation for Cancer Research (grant number CB-2011-02-01) and the Basic Science Research Program through the National Research Foundation of Korea, funded by the Ministry of Education, Science and Technology (grant number 2011-0022379), supported the study Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		

Lee 2012 (*Continued*)

			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Li 2010

Study characteristics	
Patient sampling	Retrospective consecutive? patient series.
Patient characteristics and setting	158 participants, median age = 58.4 years (range = 38 to 76), 97 males/61 females, China **Histology of primary tumour** Adenocarcinoma: N = 74; squamous cell: N = 67; adenosquamous carcinoma: N = 3; large cell carcinoma: N = 7; NOS N = 7; comorbidities: not reported **Inclusion criteria** All participants with NSCLC, who received radical surgery/mediastinoscopy for lung cancer and a preoperative FDG-PET/CT in the authors' hospital between August 2005 and December 2009 **Exclusion criteria** None listed, but it is stated that none of the participants had received any complementary therapy, such as chemotherapy or radiotherapy, before surgery or mediastinoscopy **Previous/all tests** All participants received FDG PET-CT, blood routine test, blood biochemistry test, electrocardiogram, X-ray, contrast-enhanced chest CT, and pulmonary function tests. Some participants also

	received whole body bone scan and brain MRI based on clinical needs to investigate potential metastatic lesions **Clinical setting** Secondary/tertiary setting It is unclear if the inclusion criteria narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease on PET-CT
Index tests	FDG PET-CT were acquired on a Siemens Biograph HR 16 type. The participants fasted for 4 to 6 hours, so their levels of blood glucose were controlled at a normal level. Intravenous injection of imaging FDG 0.15 to 0.2mCi/kg was undertaken. Participants were asked to lie motionless for 40 to 60 min. Whole body scans were conducted for participants after urination. PET scan parameter: slice width 4 mm, matrix 168*168, acquisition time: 2.5 min/bed. 3-Dimensional image reconstruction of PET-CT were conducted through Wizard workshop fusion images and MSViewer. A nuclear medicine specialist determined the SUV measurement. The size and location of focus were accurately measured by sagittal and coronal fusion images. The CT value and SUV value of area of interest were also measured, so the lymph nodes of hilus of lung and mediastinum could be distinguished from primary focus. Focus values with SUVmax > 2.5 were regarded as positive. No information about attenuation correction was reported **Covariates** Type of PET-CT scanner: Siemens Biograph HR 16 type FDG dose: 0.15 to 0.2 mCi/kg Injection-to-scan time: 40 to 60 min Attenuation correction: not reported Cut-off values for test positivity (malignancy): SUVmax > 2.5
Target condition and reference standard(s)	13/158 participants received mediastinal lymph node biopsy, and 4 of these 13 participants with no lymph node metastasis received radical resection of pulmonary carcinoma. Of the 149/158 participants who received lung cancer surgery, 13 received total pneumonectomy, 2 received *double-sleeve (translator unsure)* pulmonary artery lobectomia pulmonalis (this operation could also shape the pulmonary artery and trachea), 5 received *sleeve-like* pulmonary artery lobectomia pulmonalis, 3 received bilobectomy, and 126 received pulmonary lobectomy. All surgery included systematic mediastinal lymph node dissection
Flow and timing	All participants were accounted for in the results. All participants underwent surgical staging. There were no uninterpretable results
Comparative	
Notes	The Chinese Department of Health (W2009BX027) funded the study. This study is published in Chinese, and the methods and results sections were translated by a Chinese colleague of MSH and EC. This study was therefore only data extracted and appraised by MSH who double-checked the extractions and entries Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		

Were all patients included in the analysis?	Yes		

Li 2012a

Study characteristics	
Patient sampling	Retrospective? consecutive?patient series
Patient characteristics and setting	80 patients, mean age = 58 (range = 35-84) years, 52 males/28 females, China Histology of primary tumour: Not reported. Comorbidities: Not reported Inclusion criteria: Patients with NSCLC who underwent curative surgical resection with regional lymph node dissection after 18F-FDG PET/CT and breath-hold spiral CT between October 2006 and March 2009 Exclusion criteria: None listed. **Previous/all tests** : Not reported. Clinical setting: Secondary/tertiary setting The inclusion criteria narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease on PET-CT
Index tests	FDG PET-CT were acquired on an integrated PET-CT device (GE Discovery ST16) that consisted of a PET scanner and a 16-slice CT scanner. The participants fasted for $\geq$ 6 hours, and after a normal blood glucose level was ensured ($\leq$ 8 mmol/L), received an intravenous injection of FDG 3. 7 to 4.44 MBq/kg and then rested for 50 to 70 min before undergoing the scan. 2 chest radiologists with PET-CT diagnostic experience evaluated the PET-CT datasets. Nodal status was determined by SUV associated with CT attenuation. Additionally, if FDG uptake was positive, lymph nodes with calcification or higher attenuation than the aorta on the CT images were considered benign **Covariates** Type of PET-CT scanner: integrated PET-CT device (GE Discovery ST16) FDG dose: 3.7 to 4.44 MBq/kg Injection-to-scan time: 50 to 70 min Attenuation correction: yes Cut-off values for test positivity (malignancy): Nodal positive FDG uptake was defined visually as a level greater than that of the surrounding mediastinum background
Target condition and reference standard(s)	Pathology from lung resection with mediastinal lymph node dissection
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	

Notes	The study was supported by the National Science and Technology Major Projects (2009ZX09501-026), Doctoral Fund of Ministry of Education of China (20070023041) and Beijing Hope Run Special Fund (grant LC2007A02) Adverse events: None reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low

Li 2012a (*Continued*)

DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Morikawa 2009

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	93 participants, mean age = 66.1 (SD = 10.9) years, 76 males/17 females, Japan **Histology of primary tumour** Adenocarcinoma: N = 39; squamous cell: N = 28; NSCLC NOS: N = 3; adenosquamous cell carcinoma: N = 1; small cell carcinoma: N = 4; malignant lymphoma: N = 3; melanoma (mediastinal lymph node involvement): N = 1; benign: N = 14 (sarcoidosis: N = 11, interstitial pneumonitis: N = 2, pneumoconiosis: N = 1); comorbidities: none reported **Inclusion criteria** Participants with known or suspected lung cancer and mediastinal and hilar lymph node swelling detected by chest CT. Mediastinal and hilar lymph nodes were assessed if the short axis diameter on transaxial chest CT images > 10 mm **Exclusion criteria** Participants with blood glucose levels > 126 mg/dL at the time of the FDG injection **Previous tests** None reported apart from chest CT **Clinical setting** Secondary/tertiary care
Index tests	FDG PET-CT examinations were performed with a whole-body scanner (Discovery LS; GE Healthcare). All participants fasted overnight (minimum 12 hours) before radiotracer administration. FDG PET images from the skull through the mid thigh were obtained 50 minutes after injection of 185 MBq FDG and CT-based attenuation correction was performed. CT-based attenuation correction factors were then applied to the emission data, and the attenuation-corrected emission images were reconstructed using an ordered-subset expectation maximisation iterative reconstruction algorithm. The reconstructed images were converted to SUV images using patient body weight and a dose of FDG (tumour activity concentration/injected dose/body weight). An experienced radiologist and nuclear medicine physician without knowledge of histopathologic or other radiologic data independently and prospectively interpreted FDG images . Semiquantitative analysis of the FDG uptake was based on region-of-interest analysis that produced maximal SUV and mean SUV. Swollen lymph

	nodes were evaluated using FDG PET-CT. The same radiologist and experienced nuclear medicine physician drew a region of interest over each mediastinal or hilar lymph node at the most active site on FDG PET-CT images. The optimal thresholds of maximal and mean SUV were determined using receiver operating characteristics curve-based analysis, and a maximal SUV of 4.1 and a mean SUV of 3.5 were adopted as optimal cut-off values for analysis **Covariates** Type of PET-CT scanner: integrated PET-CT scanner (Discovery LS; GE Healthcare) FDG dose: 185 MBq Injection-to-scan time: 50 min Attenuation correction: yes Cut-off values for test positivity (malignancy): maximal SUV of 4.1 and a mean SUV of 3.5. These were not prespecified
Target condition and reference standard(s)	Surgical resection, mediastinoscopy, or TBNA (surgical confirmation always obtained in cases of negative TBNA). 137 lymph nodes were studied (in the 93 participants), of which 82 were malignant, and 55 were benign. 19 malignant and 37 benign lymph nodes were diagnosed using surgery or mediastinoscopy, and 63 malignant and 18 benign lymph nodes were diagnosed using TBNA (no corresponding participant-based information reported)
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	The 21st Century COE program 'Biomedical Imaging Technology Integration Program' funded by the Japan Society for the Promotion of Science supported the study Adverse events: none reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			

Morikawa 2009 (*Continued*)

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Ohnishi 2011

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	120 participants, median age = 69 (range = 40 to 85) years, 79 males/41 females, Japan **Histology of primary tumour** (only available for 84/120 participants) Adenocarcinoma: N = 47; squamous cell: N = 19; adenosquamous cell carcinoma: N = 2; bronchioalveolar carcinoma: N = 5; large cell carcinoma: N = 2; mucoepidermoid carcinoma: N = 1;

	large cell neuroendocrine carcinoma: N = 3; benign: N = 5 (sarcoidosis: N = 1, atypical adenomatous hyperplasia: N = 2, intrapulmonary lymph node: N = 1, pulmonary tuberculosis: N = 1); comorbidities: diabetes (N = 8); no others reported **Inclusion criteria** Participants with newly diagnosed or suspected lung cancer based on CT findings whose clinical TNM stage was < T4, any N and M0 based on CT. All participants with potentially resectable lung cancer were included; therefore, both participants with and without swollen mediastinal lymph nodes regardless of location were included **Exclusion criteria** Participants with poor medical conditions of grades 4 and 5 according to the American Society of Anesthesiologists Physical status classification system and participants with bleeding tendency and coagulopathy **Previous tests** None reported apart from chest and upper abdominal CT **Clinical setting** Secondary/tertiary care
Index tests	PET-CT was performed using a Discovery ST Elite Performance scanner (GE Healthcare, Tokyo, Japan) with whole-body attenuation. After participants had fasted for 4 hours and their blood glucose levels had been confirmed to be < 200 mg/dl, they were injected with an intravenous dose of 190 to 300 MBq FDG. The emission study commenced 50 min later. FDG uptake in the mediastinum was first examined based on visual interpretation. Lymph nodes with significantly higher accumulation of FDG than surrounding mediastinum levels were identified, and a SUV > 3 was judged to be positive **Covariates** Type of PET-CT scanner: Discovery ST Elite Performance scanner (GE Healthcare, Tokyo, Japan) FDG dose: 190 to 300 MBq Injection-to-scan time: 50 min Attenuation correction: yes Cut-off values for test positivity (malignancy): SUV > 3 was judged to be positive
Target condition and reference standard(s)	Pathological results from surgery, EBUS-TBNA, and EUS-FNA
Flow and timing	Data only available from 110/120 participants
Comparative	
Notes	No details of funding were reported, but the authors report that they had no competing interests Adverse events: none reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Ohno 2011

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	250 participants, mean age = 73 (range = 61 to 83) years, 136 males/114 females, Japan **Histology of primary tumour** Adenocarcinoma: N = 218; squamous cell: N = 23; adenosquamous cell carcinoma: N = 3; large cell carcinoma: N = 6; comorbidities: not reported **Inclusion criteria** Quote: "Between January 2009 and December 2010, 250 consecutive patients with T1 or T2 NSCLC as evaluated on chest radiographs or CT images at a nearby hospital underwent contrast material-enhanced multidetector CT, STIR turbo SE imaging, DW MR imaging, FDG-PET/CT, and mediastinoscopy before thoracotomy or resection of the primary lesion or before thoracotomy for primary resection in conjunction with hilar and mediastinal sampling. All patients were followed up for more than 1 year." No other information reported **Exclusion criteria** None listed **Previous tests** None listed apart form those above **Clinical setting** Secondary/tertiary care
Index tests	All participants fasted for at least 6 hours before intravenous administration of FDG at a rate of 3. 3 MBq per kilogram of body weight, and images were obtained from the skull to the mid thigh 60 minutes after completion of injection. All FDG PET-CT examinations were performed with a commercially available PET-CT scanner (Discovery ST; GE Healthcare, Milwaukee, WI, USA) . Unenhanced CT was performed from the skull to the pelvic floor according to a standardised protocol, and immediately afterwards, PET was performed in the identical transverse field of view. All integrated PET-CT examinations were performed within 60 minutes. Visual and quantitative assessment of FDG uptake was performed. To assess the validity of qualitative analysis of FDG PET-CT images, 2 general radiologists, with 8 and 14 years of experience, prospectively evaluated all FDG PET-CT images. These 2 general radiologists also had more than 3 years of experience with PET. The probability of lymph node metastasis was evaluated on a per-node basis with the following 5-point visual scoring system: 1 = definitely absent, 2 = probably absent, 3 = equivocal, 4 = probably present, and 5 = definitely present. Each reviewer assessed all lymph nodes twice. The final probability of lymph node metastasis was based on consensus of the 2 readers. For quantitative assessment of lymph node metastases on integrated FDG PET-CT images, all SUVmax measurements were obtained from regions of interest drawn over mediastinal and hilar lymph nodes (range = 16 to 576 mm^2) on the axial and coronal planes by the same chest radiologist and averaged to determine the final value for each lymph node. In addition, each investigator who performed quantitative or qualitative assessment with all modalities (see Inclusion criteria) had no knowledge of the results of pathologic examination or any other investigation and evaluated all examinations in random order **Covariates** Type of PET-CT scanner: integrated PET-CT scanner (Discovery ST; GE Medical systems) FDG dose: 3.3 MBq/kg Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): The feasible threshold value of each method was determined as demonstration of the highest accuracy, sensitivity, or both, and this MAY be SUVmax

Ohno 2011 (*Continued*)

	≥ 4 for qualitative assessment and SUVmax ≥ 2 for quantitative assessment (Tables 1 and 2 and Figures 3 and 4). These cut-off values were not prespecified
Target condition and reference standard(s)	Pathologic findings in resected specimens. Cervical mediastinoscopy (supplemented by left anterior mediastinotomy if the tumour was in the left upper lobe) if CT findings suggested invasion of the superior mediastinum or enlarged nodes plus complete standard mediastinal nodal sampling (N = 23) or thoracotomy plus systematic ipsilateral hilar and mediastinal sampling at the locations specified by the regional lymph node classifications of the American Joint Committee on Cancer and the Union for International Cancer Control (N = 227)
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	Philips Healthcare and a grant-in-aid for scientific research from the Japanese Ministry of Education, Culture, Sports, Science and Technology provided the study with financial, technical support, or both. None of the authors was an employee of Philips Healthcare, and the authors had full control over the data for the duration of this study 2 2-by-2s (qualitative and quantitative measurement): qualitative: TP = 14, FP = 0, FN = 7, TN = 229 quantitative: TP = 16, FP = 0, FN = 5, TN = 229 Adverse events: none reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Ohno 2011 (*Continued*)

Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Ose 2012

Study characteristics	
Patient sampling	Retrospective consecutive? patient series
Patient characteristics and setting	112 participants, mean age = 65.6 (SD = 9.7; range 31 to 86) years, 71 males/41 females, Japan **Histology of primary tumour** Adenocarcinoma: N = 84; squamous cell carcinoma: N = 16; large cell carcinoma: N = 4; other: N = 8; comorbidities: not reported **Inclusion criteria** Lung cancer participants who between March 2006 and February 2010 underwent preoperative examinations with PET-CT followed by radical resection with hilar and mediastinal dissection to confirm the histopathological diagnosis of lymph node metastasis at the authors' institution **Exclusion criteria**

Ose 2012 (*Continued*)

	Not reported **Previous tests** Contrast-enhanced chest CT. No further tests reported **Clinical setting** Department of General Thoracic Surgery
Index tests	**Covariates** Type of PET-CT scanner: whole-body scanner. No further information reported FDG dose: 3.7 MBq/kg Injection-to-scan time: 60 min Attenuation correction: not reported Cut-off values for test positivity (malignancy): Lymph nodes were diagnosed as metastatic when SUVmax > 2.5 or short axis > 1 cm on CT
Target condition and reference standard(s)	Pathology from radical resection and lymph node dissection
Flow and timing	Data missing for 3/112 from Table 2, but according to Table 3, the sensitivity for N2 is 50%, which means that the missing data must be TPs
Comparative	
Notes	No details of funding were reported, but the authors declared "no conflicts of interest" Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		

Ose 2012 (*Continued*)

Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Ozkan 2011

Study characteristics	
Patient sampling	Retrospective patient series
Patient characteristics and setting	153 NSCLC participants, mean age = 61.4 (SD = 9.97) years, 135 males/18 females, Turkey **Histology of primary tumour** Not reported; comorbidities: not reported **Inclusion criteria** Not reported **Exclusion criteria** Not reported **Previous tests** Not reported

Ozkan 2011 (*Continued*)

	Clinical setting Secondary/tertiary care
Index tests	FDG-PET-CT **Covariates** Type of PET-CT scanner: not reported FDG dose: not reported Injection-to-scan time: not reported Attenuation correction: not reported Cut-off values for test positivity (malignancy): not reported
Target condition and reference standard(s)	Histopathology or clinical follow-up decisions and the results of other imaging tests
Flow and timing	If the whole sample only consisted of the 153 participants reported in the abstract, all participants were included in the analysis
Comparative	
Notes	Author sent test accuracy data on request. As the study was only published as an abstract, we contacted the author (on 3 December 2012) to request the following information How was the sample recruited? - Did the study population consist of a consecutive sample? Characteristics of the 153 participants: - Histology of primary tumour? - Comorbidities? - NSCLC stage? - Inclusion criteria? - Exclusion criteria? - Previous/all reported tests? - Clinical setting Thoracic surgery unit? - Were any participants excluded from the analyses? PET-CT scanning: - What type of PET-CT scanner was used? - What FDG dose was used? - What was the injection-to-scan time? - Did you use attenuation correction? - What was the criteria for a positive result? - Did you use a prespecified cut-off value for test positivity? - Was the PET-CT results interpreted without knowledge of the pathological/follow-up results? Pathological staging: You state in the abstract that PET/CT results were confirmed histopathologically or according to clinical follow-up decisions and the results of other imaging tests - How many participants received histopathological confirmation, and how many participants received non-histopathological confirmation and why (including, what did their confirmation consist of instead)? - Was the pathological staging results interpreted without knowledge of the PET-CT results? Flow and timing:

	- What was the interval between PET-CT and histopathological/other confirmation of the PET-CT results? - Were there any adverse events of the PET-CT? Funding - Was the study funded and if yes, by whom? We received no response Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	Unclear		
Was a positive result defined?	No		
			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Ozkan 2011 (*Continued*)

			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		

Perigaud 2009

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	51 participants, mean age = 60.6 (SD = 9.3) years, 44 males/7 females, France **Histology of primary tumour** Adenocarcinoma: N = 23; squamous cell: N = 24; large cell: N = 4; comorbidities: not reported **Inclusion criteria** Consecutive participants with suspected or pathologically proven non-small cell lung cancer from June 2006 to February 2008. All these participants had previously undergone a classic staging procedure and an integrated FDG PET-CT in the authors' centre, and a decision to perform first-line surgery was taken on the basis of the FDG PET-CT. (Mediastinoscopy was performed in 2 participants to exclude contralateral mediastinal lymph node involvement (N3)) **Exclusion criteria** None listed **Previous tests** The preoperative assessment also comprised classic staging (bronchoscopy and chest, abdominal, and brain CT scan) and functional assessment comprising pulmonary function tests, arterial blood gases, and possibly, pulmonary perfusion scintigraphy **Clinical setting** Secondary/tertiary care The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	Integrated FDG PET-CT was performed with a Discovery LS PET-CT system (General Electric Medical Systems, Milwaukee, WI, USA), consisting of an Advance NXi PET scanner and a 4-slice Light Speed Plus CT. Participants fasted for 6 hours prior to undergoing the FDG-PET/CT acquisitions, and blood glucose was required to be less than 7 mmol/l before the intravenous injection of 310 to 45 MBq (5 to 7 MBq/kg) of FDG. FDG injection was followed by a period

	of approximately 60 min. Participants were allowed to breathe normally during PET and CT acquisitions. Unenhanced chest CT image was obtained from the participant's integrated FDG PET-CT with the use of a standardised protocol involving 140 kV, 90 mA, a tube rotation time of 0. 5 s per rotation, a pitch of 6, and a section thickness of 5 mm. PET images were reconstructed using CT data, for attenuation correction, using the ordered subsets expectation maximisation algorithm and without CT-based attenuation correction. A nuclear radiology physician and a thoracic surgeon blinded to the PET data interpreted CT data. The short axis diameters of the largest lymph nodes in each mediastinal station harvested during surgery were measured on CT images. Lymph nodes not visualised by CT were considered to be non-measurable. A nuclear radiology physician interpreted the PET-CT data. The regions of interest (ROI) were placed manually over the areas of the lung tumour mass and the mediastinal node(s). The SUVmax was calculated based on the measured activity, decay-corrected injected dose, and the participant's body weight **Covariates** Type of PET-CT scanner: integrated PET-CT scanner (Discovery LS PET-CT system, General Electric Medical Systems, Milwaukee, WI, USA) FDG dose: 5 to 7 MBq/kg Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): An arbitrary SUVmax cut-off of 3 was used for mediastinal lymph nodes
Target condition and reference standard(s)	All participants were treated by pulmonary resection. Most operations were performed via a postero-lateral thoracotomy. A left pneumonectomy was performed via sternotomy with cardiopulmonary bypass in 1 case because of tumour invasion of the bifurcation of the pulmonary artery. Right pneumonectomy was associated with a carinal resection in 1 case. Sleeve lobectomies were performed in 3 participants. Extensive mediastinal lymph node dissection was performed systematically, with resection of stations 2R, 3a, 4R, 7, 8, and 9 on the right side, and stations 5, 6, 7, 8, and 9 on the left side when these lymph nodes were present. All lymph nodes removed were embedded whole and examined by an experienced histologist. Immunohistochemistry was performed at the histologist's discretion. The mean interval between PET-CT and surgery was 31 (SD = 15.8) days
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	No details of funding were reported. However, the study was probably not externally funded because although this data collection was prospective, it appears to be collected as part of normal practice 2 of the 51 participants had received chemotherapy, and the last course had been administered more than 1 month prior to the FDG-PET/CT imaging Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Perigaud 2009 (*Continued*)

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Plathow 2008

Study characteristics	
Patient sampling	Prospective patient series
Patient characteristics and setting	52 participants, mean age = 62 (range = 49 to 71) years, 36 males/16 females, Germany **Histology of primary tumour** Not reported; comorbidities: not reported **Inclusion criteria** Participants with histologically proven NSCLC who were referred to the authors' department for staging by PET-CT and whole-body MRI. According to staging diagnostic procedures performed elsewhere, all the participants had stage IIIa (N = 4 participants) or IIIb (N = 48). The participants were sent to the authors' department to check for potential operability **Exclusion criteria** None listed **All/Previous tests** See **Inclusion criteria** **Clinical setting** Departments of Radiology/Nuclear Medicine
Index tests	All participants fasted overnight to guarantee low blood sugar levels. PET-CT scanning (from the base of the skull to the lower legs) started 55 to 65 minutes after intravenous administration of 370 MBq (mean = 370 MBq, range = 360 to 400 MBq) of FDG and was performed with Hi-Rez Biograph 16 (Siemens Medical Solutions, Knoxville, TN, USA) consisting of a high-resolution 3D LSO PET and a state-of-the-art 16 row multislice CT. In all participants, a multiphase CT protocol with an intravenous application of 120 ml iodinated contrast agent (Ultravist 370, Schering GmbH Berlin, Germany) was performed. 2 specialists in nuclear medicine who were aware of the clinical status of the participant, but were blinded to the results of the other imaging studies and previous tests, independently interpreted the PET-CT data **Covariates** Type of PET-CT scanner: Hi-Rez Biograph 16 (Siemens Medical Solutions, Knoxville, TN, USA) FDG dose: 370 MBq Injection-to-scan time: 55 to 65 min Attenuation correction: yes Cut-off values for test positivity (malignancy): The FDG-distribution was rated visually and optionally quantified as SUVs. Any focal tracer uptake exceeding normal regional tracer accumulation was assessed as a malignant lesion. The determination of malignancy on CT was based on morphological characteristics (e.g., invasive and irregular growth pattern) and enhancement pattern. Lymph node involvement on CT was based on region specific nodal size criteria based on measurement of the small axis diameter
Target condition and reference standard(s)	Mediastinoscopy ± staging obtained by surgery

Plathow 2008 (*Continued*)

Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results		
Comparative			
Notes	During January and February 2011, we tried to contact the corresponding author a variety of ways via email to ascertain whether this paper dealt with patient staging (as stated in the Materials and Methods section) or restaging (as stated in the abstract), but it is unclear whether any of our emails reached the corresponding author, and no response was received Source of funding: no details reported Adverse events: not reported		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Plathow 2008 (*Continued*)

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Sanli 2009

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	78 participants, mean age = 61.3 (range = 44 to 79) years, 73 males/5 females, Turkey **Histology of primary tumour** Adenocarcinoma: N = 16; epidermoid carcinoma: N = 41; adenosquamous carcinoma: N = 5; large cell or undifferentiated carcinoma: N = 16: comorbidities: not reported **Inclusion criteria** Consecutive participants with NSCLC who were potential candidates for surgical resection and were admitted to the thoracic surgery unit of the authors' hospital from March 2006 to June 2008 **Exclusion criteria** Participants with evidence of metastatic disease, except for those with solitary brain or adrenal metastasis, participants who had not undergone PET-CT scanning as part of their preoperative evaluation or who had undergone FDG PET scanning in another centre, participants with diabetes mellitus whose blood glucose levels could not be controlled and brought to normal values, and participants receiving neoadjuvant treatment **All/Previous tests** Complete blood counts and blood chemistry tests, chest radiographs, thoracic CT scans, PET-CT scans, pulmonary function tests, and if clinically indicated, bone scans and cranial magnetic resonance imaging were performed in all cases **Clinical setting** Thoracic surgery unit It is unclear if the inclusion criteria narrow the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a

	proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	Whole-body PET-CT scanning was performed with Siemens Biograph 2 PET-CT system (Siemens, Munich, Germany). A whole-body acquisition was performed immediately 1 hour after intravenous administration of FDG (11 to 16 mCi), and images were obtained from the vertex to the upper thigh region. High-quality images were acquired, and semiquantitative measurements of glucose metabolism were obtained. All participants fasted for at least 4 hours before imaging; their fasting blood glucose levels were within the normal range, and none received insulin to return blood glucose to normal levels. The SUVs of hilar lymph nodes and MLNs were determined from the transverse views by the nuclear medicine physician blinded to results of reference tests. Coronal-sagittal images and their correlation with CT scans were used when the exact location was uncertain Regions of interest were drawn on the images, and semiquantitative SUV measurements were defined as the regional tissue radioactivity concentration normalised for injected dose and body weight. There were no details reported about attenuation correction **Covariates** Type of PET-CT scanner: Siemens Biograph 2 PET-CT system (Siemens, Munich, Germany) FDG dose: 11 to 16 mCi Injection-to-scan time: 60 min Attenuation correction: not reported Cut-off values for test positivity (malignancy): Results of PET-CT scans were considered positive in the mediastinum and hilar area that was separate from the primary mass if the SUV in participants suspected to have lymph node metastases was > 2.5. Please note, it is not specified that it is SUVmax; however, for the subgroup analyses of the criteria for test positivity, we categorised this study in the *SUVmax ≥ 2.5* group
Target condition and reference standard(s)	Definitive diagnosis was established based on the histopathologic findings of lymph node sampling in mediastinoscopy or biopsy during the surgical procedure. If a mediastinoscopy was performed, histologic evaluation of the specimens was performed with a frozen section. If N2 disease was present, definitive resection was not performed at that time. These participants received neoadjuvant therapy. Participants with resectable disease on mediastinoscopy underwent further operative procedures. 3 participants underwent transcarinal sleeve pneumonectomy, 16 underwent pneumonectomy (invasion of the left atrium, main pulmonary artery, carina, distal trachea, and proximal main bronchus, with some major fissure invasion), 4 underwent bilobectomy, 2 underwent sleeve lobectomy, and 46 underwent lobectomy. 1 participant was identified as unresectable (M1) during thoracotomy. 6 participants did not undergo thoracotomy because of positive results on mediastinoscopy. 1 participant given a positive diagnosis after mediastinoscopy underwent resection as a result of drainage to the pleural space caused by tumour necrosis and haemoptysis. Multistation nodal mediastinal sampling was performed, with removal of levels 2, 4, 7, 8, and 9 on the right side. For left-sided tumours, lymph nodes at levels 5 and 6 were dissected also. However, non-palpable station 2L could not be removed in some participants. Hilar lymph nodes were also dissected. All participants underwent tissue sampling of MLNs to compare sampling results with imaging results
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	

Notes	No details of funding were reported. However, the study was probably not externally funded because although this data collection was prospective, it appears to be collected as part of normal practice Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
			Low
DOMAIN 4: Flow and Timing			

Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Saydam 2012

Study characteristics	
Patient sampling	Retrospective patient series
Patient characteristics and setting	42 participants, mean age = 67 (range = 56 to 78) years, 42 males/0 females, Turkey **Histology of primary tumour** Adenocarcinoma: N = 14; squamous cell carcinoma: N = 27; large cell carcinoma: N = 1; comorbidities: not reported **Inclusion criteria** Quote: "This study was performed in between June 2007 and June 2010, in which [the authors] retrospectively reviewed 42 (42 men) retired coal workers who had NSCLC with no distant metastasis" **Exclusion criteria** None listed **All/Previous tests** Chest X-ray, contrast-enhanced thoracic CT scans, bronchoscopy or transthoracic needle aspiration biopsy, whole-body PET-CT scans, and cranial MRI **Clinical setting** Departments of Thoracic Surgery and Medical Oncology
Index tests	Participants fasted for 4 to 6 hours prior to the FDG injection, and blood glucose was measured before this injection to ensure levels < 200 mg/dl **Covariates** Type of PET-CT scanner: not reported FDG dose: 370-555 MBq (10 to 15 mCi) Injection-to-scan time: 60 min Attenuation correction: not reported Cut-off values for test positivity (malignancy): SUV > 2.5 in mediastinal lymph nodes > 1 cm. Please note, it is not specified that it is SUVmax; however, for the subgroup analyses of the criteria for test positivity, we categorised this study in the *SUVmax ≥ 2.5* group
Target condition and reference standard(s)	Cervical or extended mediastinoscopy with or without thoracotomy including mediastinal node sampling

Flow and timing	All participants were accounted for. All participants received the reference standard, and there were no uninterpretable results		
Comparative			
Notes	The authors report that this study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors Adverse events: none reported		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge	Unclear		

of the results of the index tests?			
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Shin 2008

Study characteristics	
Patient sampling	Retrospective consecutive? patient series
Patient characteristics and setting	184 participants, mean age = 59 (SD = 10, range = 32 to 81) years, 124 males/60 females, South Korea **Histology of primary tumour** Adenocarcinoma: N = 132; squamous cell carcinoma: N = 40; large cell neuroendocrine carcinoma: N = 5; large cell carcinoma: N = 4; sarcomatoid carcinoma: N = 2; pleomorphic carcinoma: N = 1; comorbidities: not reported **Inclusion criteria** Participants who underwent integrated PET-CT and surgical staging from September 2003 to July 2006 and were diagnosed with stage T1 NSCLC at CT (retrospective review and reanalysis of all CT data obtained during this period) or on pathologic examination. It is unclear if it is all participants within the time period **Exclusion criteria** Participants with bronchioloalveolar carcinomas (nodule of pure ground-glass opacity) because these usually do not have nodal or extra-thoracic metastases **Previous tests** Not explicitly stated, but it is implied that participants underwent conventional staging (clinical examination or enhanced thoracic CT covering down to the level of middle portion of the kidneys) **Clinical setting** Secondary/tertiary
Index tests	Peripheral blood glucose was < 150 mg/dL in all participants. Participants received an intravenous injection of 370 MBq (10 mCi) of FDG and then rested for over 45 minutes before scanning. Scans were acquired using a PET-CT unit (Discovery LS; GE Healthcare, Milwaukee, WI, USA) consisting of a PET scanner (Advance NXi; GE Healthcare) and an 8-slice CT scanner (Light Speed Plus; GE Healthcare). Immediately after an unenhanced CT, an emission PET was performed in the

	identical transverse field of view. A chest radiologist (with 18 years of CT interpretation experience and 4 years of PET-CT interpretation) and a nuclear medicine physician (with 13 years of experience and 4 years of PET-CT analysis) jointly and prospectively evaluated integrated PET-CT images. Both were unaware of the findings of clinical and pathological evaluations **Covariates** Type of PET-CT scanner: Discovery LS (GE Healthcare, Milwaukee, WI, USA) FDG dose: 370 MBq (10 mCi) Injection-to-scan time: 45 min Attenuation correction: yes Cut-off values for test positivity (malignancy): All lymph nodes in the thorax with abnormal FDG uptake (> mediastinal blood pool uptake), irrespective of their size, were considered metastatic. Enlarged lymph nodes with their short axis diameter > 10 mm were designated as benign when they were negative at the PET component of PET-CT images. Additionally, lymph nodes even with high FDG uptake, when they showed higher attenuation than mediastinal structures (great vessels) or benign calcification (central, nodular, diffuse, laminated, or popcorn-like), were regarded as being benign
Target condition and reference standard(s)	Surgical staging (mediastinoscopy with routine node sampling of areas 2R, 4R, 2L, 4L, and 7 / thoracotomy with sampling of all visible and palpable lymph nodes that were accessible in the hilum and mediastinum (all encountered lymph nodes were removed from the ATS lymph node map areas 10R, 9, 8, 7, 4R, 3 and 2R in tumours of the right lung, and from map areas 10L, 9, 8, 7, 6, 5, and 4L of the left lung), and when necessary, especially when imaging results suggested the presence of possible nodal metastasis in nodal stations of group 1 (highest mediastinal) or 2L (when tumour was located in the left lung) nodes, the nodes were also evaluated during a mediastinoscopy or thoracotomy)
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results. Originally, 199 participants were enrolled; of these, 15 were excluded because of not undergoing surgical treatment (N = 10) or death caused by an unrelated disease (postoperative adult respiratory distress: N = 2, chronic empyema in the lobectomy space and its related pneumonia and sepsis: N = 1, advanced gastric cancer: N = 1, and acute pulmonary thromboembolism: N = 1) during the follow-up period
Comparative	
Notes	The study was supported by SRC/ERC programme of the MOST/KOSEF (R11-2002-103) Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		

Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Sommer 2012

Study characteristics	
Patient sampling	Prospective consecutive? patient series
Patient characteristics and setting	33 participants, mean age = 63.7, median age = 66 (range = 43 to 83) years, 24 males/9 females, Switzerland **Histology of primary tumour** Adenocarcinoma: N = 16; squamous cell carcinoma: N = 8; adenosquamous carcinoma: N = 1, large cell carcinoma: N = 2, well-differentiated neuroendocrine carcinoma: N = 1; non-pulmonary primary malignancy: N = 2, benign: N = 3; comorbidities: not reported **Inclusion criteria** Participants with suspected NSCLC who underwent PET-CT and were scheduled for surgery according to the PET-CT findings. No further details reported **Exclusion criteria** None listed **Previous tests** Not explicitly stated, but also underwent whole-body MRI before surgery **Clinical setting** Secondary/tertiary
Index tests	**Covariates** Type of PET-CT scanner: integrated PET/CT system with 16-slice CT (Discovery; GE Healthcare, Chalfont St Giles, UK) FDG dose: 5 MBq/kg up to maximum dose of 500 MBq Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): Initial assessment of the PET-CT examinations was done by 2 board-certified nuclear medicine physicians with > 5 and > 10 years experience, respectively. Both low-dose and contrast-enhanced CT datasets were used for the PET-CT readings, and previous images from investigations other than PET-CT and MRI were available to all readers. The PET-CT images were interpreted in a qualitative manner considering both morphological and functional information. Increased FDG uptake was identified by visual comparison of the lesion's signal to the FDG uptake of the liver parenchyma. Quantitative values for SUV were calculated and used for interpretation of particular findings if found appropriate by the readers; however, no general cut-off values were applied for differentiating benign from malignant lesions. For lymph node assessment, a short axis diameter > 1 cm was regarded as a morphological criterion for metastatic involvement. Lymph node stations were divided into 3 groups (N1, N2, and N3), and a group of lymph node stations was rated positive if at least 1 lymph node from 1 of the stations was considered to be metastatic. The results are presented as the average values from the 2 readers
Target condition and reference standard(s)	Histological or cytological specimens (no further details provided)
Flow and timing	The 2 participants with non-pulmonary primary malignancy were excluded from the analyses
Comparative	
Notes	Funding: M Klarhöfer is an employee of Siemens Switzerland Ltd., Healthcare Sector, Zürich, Switzerland. The authors acknowledged financial support from Guerbet, Switzerland, but stated

	that the study sponsor played no role in matters of design, collection, analysis, interpretation of data, and writing the report Adverse events: none reported Please note, it seems that decimal points cannot be entered into the results, but the results of this study (presented as the average values of 2 readers) are as follows: TP = 2, FN = 4, FP = 0.5, TN = 24.5. The numbers entered into the analyses were TP = 2, FN = 4, FP = 1, TN = 24, in order to preserve the total number of participants and not inflate the accuracy estimates

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Subedi 2009

Study characteristics	
Patient sampling	Retrospective consecutive patient series
Patient characteristics and setting	161 participants, mean age = 70.2 (range = 37 to 89) years, 85 males/76 females, UK **Histology of primary tumour** Not reported; comorbidities: not reported **Inclusion criteria** Participants with known or suspected primary bronchogenic carcinoma who underwent half-body FDG PET-CT from 1 April 2006 to 31 March 2007 **Exclusion criteria** None listed **Previous tests** Chest radiograph, CT scan of the thorax and upper abdomen, fibreoptic bronchoscopy, or image-guided transthoracic fine-needle biopsy **Clinical setting** Secondary
Index tests	A standard dose of 375 MBq of -FDG was administered intravenously. PET and CT images were acquired from skull base to upper thigh after an uptake period of 60 min on either a Discovery ST or STE PET-CT camera (GE Healthcare, Milwalkee, WI, USA). The CT component of the PET-CT was performed according to a standardised protocol with the following settings: 140 kV; 80 mA; tube rotation time, 0.5 s per rotation; pitch, 6; section thickness, 3.75 mm (to match the PET section thickness). Participants maintained normal shallow respiration during the CT acquisition. No iodinated contrast material was administered. PET-CT was regarded as negative if there was no or very low metabolic activity (below mediastinal blood pool activity) within lesions. A maximum standardised uptake value (SUV) of 2.5 was used as an arbitrary cut-off **Covariates** Type of PET-CT scanner: integrated PET-CT scanner (Discovery ST; GE Medical systems)

	FDG dose: 375 MBq Injection-to-scan time: 60 min Attenuation correction: not reported Cut-off values for test positivity (malignancy): SUVmax of 2.5
Target condition and reference standard(s)	Pathological staging: 47 underwent complete lymphadenectomy, 22 underwent systematic lymph node sampling, and 4 had no nodal tissue present in their histopathological specimen
Flow and timing	Only data from 91/161 included participants were presented. It appears that the remainder did not receive pathological confirmation of their test result
Comparative	
Notes	There was no mention of funding source, but since this was a retrospective database study, it is likely that the study received no explicit funding Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			

Subedi 2009 (*Continued*)

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	No		

Tasci 2010

Study characteristics	
Patient sampling	Retrospective consecutive patient series
Patient characteristics and setting	127 participants, mean age = 58.2 (range (?) = 41-82) years, 98 males/29 females, Turkey **Histology of primary tumour** Adenocarcinoma: N = 54; squamous cell: N = 69; large cell: N = 4; comorbidities: not reported **Inclusion criteria** All participants diagnosed with NSCLC between October 2005 and January 2007 **Exclusion criteria** 65 mg/dl < fasting blood glucose > 150 mg/dl, type I diabetes mellitus, malign pleural effusion, > 30 days between PET-CT and operation date, M1 or N3 disease, neoadjuvant chemotherapy or radiotherapy, or T4 tumour with mediastinal invasion **All/Previous tests** Standard preoperative tests including past medical history, physical examination, blood tests, and spirometry. Radiological tests were done by performing chest radiography, chest computed tomography, and PET-CT **Clinical setting** Thoracic surgery unit

Index tests	Participants were imaged on a Biograph (Siemens/CTI) scanner, which produces transaxial, coronal, and sagittal reconstructions of CT, PET, and fusion PET/CT data and combines a dual detector spiral CT scanner (Somatom emotion) and a high resolution PET scanner with 4.5 mm spatial resolution and 3-dimensional image acquisition. A multimodality computer platform (Syngo; Siemens) was used for image review and manipulation. The participants had fasted for at least 6 h and then received ca. 555 MBq (15 mCi) of FDG by intravenous injection. The CT data were used for attenuation correction of PET emission images and for anatomic localisation of emission data. All PET-CT images included in the study were taken in the same centre and assessed by the same nuclear medicine expert **Covariates** Type of PET-CT scanner: Biograph (Siemens/CTI) FDG dose: ca. 555 MBq (15 mCi) Injection-to-scan time: not reported Attenuation correction: yes Cut-off values for test positivity (malignancy): The positive nodal involvement was considered to be the maximum standard uptake value (SUV) of the lymph max node that was higher than the mediastinal blood pool SUV
Target condition and reference standard(s)	Thoracotomy without invasive staging was performed in 91/127 participants who were negative for nodal spread on PET-CT. 2 participants were inoperable because of N2 nodal involvement. Invasive staging was performed with mediastinoscopy in 31/36 participants with positive involvement at PET-CT. Video-assisted thoracoscopic staging was performed in 5 participants. Mediastinal nodal involvement was determined in 12/31 participants staged with mediastinoscopy, and 3/5 participants were staged with video-assisted thoracoscopy (involvement in 8, 9 stations). The following procedures were performed in 108 participants: wedge resection in 1 because of insufficient pulmonary reserve, lobectomy in 73, bilobectomy in 6, sleeve lobectomy in 7, pneumonectomy in 18, and sleeve pneumonectomy in 3. The average interval between PET-CT scanning and the operation was 19 (1 to 30) days
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	There was no mention of funding source, but since this was a retrospective database study, it is likely that the study received no explicit funding Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Toba 2010

Study characteristics	
Patient sampling	Retrospective consecutive? patient series.
Patient characteristics and setting	42 participants, (mean?) age = 68 (SD? = 9.5) years, 25 males/17 females, Japan **Histology of primary tumour** Adenocarcinoma: N = 30; squamous cell: N = 8; large cell: N = 1; small cell: N = 2; pleomorphic carcinoma: N = 1; comorbidities: not reported **Inclusion criteria** Participants with a histologic diagnosis of lung cancer who underwent a curative operation (pulmonary resection and lymph node dissection) at Tokushima University Hospital between December 2005 and April 2007 **Exclusion criteria** None listed, but none of the participants received preoperative chemotherapy or radiotherapy **All/Previous tests** All participants underwent preoperative FDG PET-CT, chest CT, chest X-ray, and brain MRI "as well as operative indication being assessed" (?) **Clinical setting** Secondary/tertiary The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	After a 6-hour fast and avoidance of strenuous work/exercise for 24 hours, the participants received an injection of FDG (3.7 MBq/kg body weight) followed by FDG-PET/CT scanning 1 hour later using an Aquiduo (Toshiba Medical Systems, Tokyo, Japan). Participants were imaged from the skull base to mid thigh. CT was performed according to the following technical parameters: detector row configuration = 16 X 1.25 mm, a helical pitch = 15, a gantry rotation speed = 0.5 s, peak voltage = 120 kVp, a tube load = 50 mA, and a slice thickness = 2 mm. An emission scan was acquired immediately following the CT scan for 2 min per bed position in 3D. Abnormal FDG uptake was defined as visually > the background activity in surrounding normal tissue excluding physiologic uptake sites. The short axis diameter of the largest lymph node at its station was also measured and considered as metastasis if > 1 cm at axial view of CT scan (not sure if this criterion solely refers to a separate multidetector CT scan or the CT aspect of the FDG PET-CT). 2 chest surgeons analysed the PET-CT scans by consensus while referring to reports made by several experienced nuclear medicine physicians. No details about attenuation correction were reported **Covariates** Type of PET-CT scanner: Aquiduo (Toshiba Medical Systems, Japan) FDG dose: 3.7 MBq/kg body weight Injection-to-scan time: 60 min Attenuation correction: not reported Cut-off values for test positivity (malignancy): Abnormal FDG uptake was defined as visually > the background activity in surrounding normal tissue excluding physiologic uptake sites
Target condition and reference standard(s)	Pulmonary resection with lymph node dissection
Flow and timing	All participants were accounted for in the results. All participants received surgical staging. There were no uninterpretable results

Comparative			
Notes	There was no mention of funding source, but since this was a retrospective database study, it is likely that the study received no explicit funding Adverse events: not reported		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low

Toba 2010 (*Continued*)

DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Tournoy 2007

Study characteristics	
Patient sampling	Prospective consecutive (?) Patient series
Patient characteristics and setting	52 participants, median age = 68 (range = 48 to 80) years, 39 males/13 females, Belgium **Histology of primary tumour** Adenocarcinoma: N = 17; squamous cell: N = 20; large cell: N = 10; adenosquamous: N = 5; comorbidities: not reported **Inclusion criteria** Consecutive participants with suspected or pathologically proven primary NSCLC were eligible if a tissue specimen from at least 1 of the intrathoracic lymph nodes was available and if they underwent an integrated FDG PET-CT scan **Exclusion criteria** None listed **Previous tests** CT **Clinical setting** Hospital department of respiratory medicine It is unclear if the inclusion criteria narrow the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	Participants fasted for at least 6 hours, after which blood glucose levels were determined to ascertain a level of < 200 mg/dl. Participants then received 4 MBq/kg FDG intravenously followed by 250 ml sodium chloride and 20 mg furosemide. Image acquisition started 60 min after injection of FDG in a relaxed supine position with the arms alongside the body using an integrated FDG PET-CT scanner (Philips Gemini FDG PET-CT, Philips Medical Systems, Cleveland, OH, USA). First, a total body low-dose CT scan for calculation of the attenuation correction was performed (120 kV, effective tube current-time product maximum 30 mAS, pitch 0.9, collimation 1661.5 mm, rotation time 0.5 s, reconstructed contiguous slices of 5 mm, scan field from head up to the upper tights). Second, a scan was performed with a dual head injector (175 mAS, otherwise the same scan parameters) after intravenous injection of 120 ml contrast medium with an iodine concentration of

	300 mg/ml at a flow rate of 1.8 ml/s followed by a saline flush. No oral contrast was administered. Next, the FDG PET scan from the orbitomeatal region up to the upper tights (consisting of 8 to 9 bed positions of 3 min per table position) was performed. The CT and integrated FDG PET-CT scans represented a single procedure of data acquisition but were read separately. For the CT analysis, the radiologist was blinded to the FDG PET data. All intrathoracic lymph nodes were noted and the small and long axes were measured (mm). A lymph node with a short axis of at least 10 mm was indicated as suspect. The FDG PET-CT scan was interpreted based on both CT and FDG PET images, which were read by a nuclear physician and a radiologist. The maximum and mean SUV values were determined by drawing regions of interest on the attenuation-corrected PET fusion images around the primary tumour or the involved lymph node. The variables SUVmax and SUVmean were then calculated as the maximum and mean SUV values, respectively, within the region of interest. Quantitative evaluation based on the SUVmax/SUVliver ratio was calculated as the ratio of the SUVmax over the mean SUV value obtained from the homogenous distribution of radioactivity in the liver. The ultimate rating of positive/negative per-patient result "is based on the visual correlation by the nuclear physician" **Covariates** Type of PET-CT scanner: integrated FDG PET-CT scanner (Philips Gemini FDG PET-CT; Philips Medical Systems, Cleveland, OH, USA) FDG dose: 4 MBq/kg FDG intravenously followed by 250 ml sodium chloride and 20 mg furosemide Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): The ultimate rating of positive/negative per-patient result "is based on the visual correlation by the nuclear physician"
Target condition and reference standard(s)	For intrathoracic lymph nodes, a tissue sample was obtained either by mediastinoscopy, surgical resection, or by linear endoscopic ultrasound. The latter consisted of either oesophageal endoscopic ultrasound with real-time guided fine needle aspiration (EUS-FNA) or endobronchial endoscopic ultrasound with real-time guided transbronchial needle aspiration (EBUS-TBNA). Because the negative predictive values of EUS-FNA and EBUS-TBNA were considered too low, surgical confirmation was always done in case no malignant lymph node invasion could be demonstrated by either of these endoscopic techniques
Flow and timing	Data were available for 48/52 participants because for 2/52 participants, no pathologically confirmed mediastinal data were available (only hilar), and for 2/52 participants, the central location of the primary tumour precluded a confident discrimination of mediastinal nodes
Comparative	
Notes	Per-patient data provided by an author via email communication The study received no funding Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Unclear		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Uruga 2011

Study characteristics	
Patient sampling	Retrospective consecutive patient sample
Patient characteristics and setting	182 participants; participants without interstitial pneumonia (only reported for 139/159): median age = 67 (range = 30 to 88) years, 77 males/62 females; participants with interstitial pneumonia (only reported for 21/23): median age = 71 (range = 54 to 88) years, 17 males/4 females; Japan **Histology of primary tumour** - Participants without interstitial pneumonia (only reported for 139/159): adenocarcinoma: N = 118; squamous cell: N = 15; large cell neuroendocrine carcinoma: N = 2; pleomorphic carcinoma; N = 0; other: N = 4 - Participants with interstitial pneumonia (only reported for 21/23): adenocarcinoma: N = 12; squamous cell: N = 6; large cell neuroendocrine carcinoma: N = 1; pleomorphic carcinoma: N = 1; other: N = 1; comorbidities: 23/182 participants had interstitial pneumonia **Inclusion criteria** Quote: "We reviewed medical records of participants in our hospital between April 2008 and July 2010. Patients were included in our study if they had undergone lobectomy and conventional lymph node dissection for NSCLC, FDG PET-CT and a contrast-enhanced CT scan within 60 days of the operation, and their blood glucose level was below 150 mg/dl before FDG PET-CT examination" **Exclusion criteria** Participants with lymphoproliferative disorders (N = 1) **Previous tests** CT **Clinical setting** Not reported The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	FDG PET-CT **Covariates** Type of PET-CT scanner: Toshiba Aquiduo 16 FDG dose: 185 MBq Injection-to-scan time: 60 to 120 minutes Attenuation correction: no Cut-off values for test positivity (malignancy): Lymph nodes with a standardised uptake value max (SUVmax) above 2.5 were considered metastatic. Obvious calcifications on CT component images of PET-CT were considered as benign
Target condition and reference standard(s)	Lobectomy and selective lymphadenectomy (conventional lymph node dissection)
Flow and timing	Data were reported for the 182 reported participants

Uruga 2011 (*Continued*)

Comparative			
Notes	Published as an abstract only. The author was therefore contacted (on 18 June 2013) to request the missing information, and a response was received (on 19 June 2013) The study received no funding Adverse events: none (PET-CT). Reference standard: Not reported Participants with interstitial pneumonia (N = 23): TP = 1, FN = 6, FP = 4, TN = 12 Participants without interstitial pneumonia (N = 159): TP = 9, FN = 14, FP = 11, TN = 125		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	Yes		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge	Yes		

Uruga 2011 (*Continued*)

of the results of the index tests?			
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Uskul 2009

Study characteristics	
Patient sampling	Retrospective patient series
Patient characteristics and setting	37 participants, mean age = 59 (SD? = 9) years, 34 males/3 females, Turkey **Histology of primary tumour** Adenocarcinoma: N = 11; squamous cell, N = 26; comorbidities: not reported **Inclusion criteria** Participants with NSCLC who underwent TBNA (i.e., who had mediastinal nodes ≥ 10 mm on CT) of mediastinal lymph nodes during fibreoptic bronchoscopy and PET-CT examination at the authors' institution during a 2-year period **Exclusion criteria** None listed **Previous tests** CT **Clinical setting** Secondary/tertiary
Index tests	PET-CT scans were performed using a multidetector CT integrated high-resolution PET-CT scanner (Siemens Biograph LSO HI-RES Integrated PET-CT Scanner; Siemens Medical Solutions, Knoxville, TN, USA) 60 min after an intravenous injection of FDG at a dose of 5 MBq/kg body weight. To minimise insulin activity, participants were required to fast for a minimum of 12 hours prior to FDG administration. The CT component of the procedure was performed without intravenous contrast and with low current (70 mA, 5 mm section thickness), and was only used for the purpose of attenuation correction and assistance in the localisation of the PET images. 2 experienced nuclear medicine physicians, who were blind to the pathology results, evaluated images, and their consensus was classified as either negative (no typical uptake for malignancy, SUV < 2.5) or positive (typical uptake for malignancy, SUV ≥ 2.5) **Covariates**

	Type of PET-CT scanner: Siemens Biograph LSO HI-RES Integrated PET-CT Scanner (Siemens Medical Solutions, Knoxville, TN, USA) FDG dose: 5 MBq/kg Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): Consensus was from 2 nuclear medicine physicians and classified as either negative (no typical uptake for malignancy, SUV < 2.5) or positive (typical uptake for malignancy, SUV ≥ 2.5). Please note, it is not specified that it is SUVmax; however, for the subgroup analyses of the criteria for test positivity, we categorised this study in the *SUVmax ≥ 2.5* group
Target condition and reference standard(s)	Pathological staging (TBNA and surgical resection). All TBNA preparations positive for malignancy were considered true positives; whereas, only those TBNA preparations negative for malignancy that were confirmed by mediastinoscopy/mediastinal lymph node dissection were considered true negatives. All inadequate TBNA samples or adequate negative TBNA samples not confirmed by surgery were considered false negatives. All the TBNA data were examined by the same pathologist who was blinded to the participants' data
Flow and timing	1/37 included participants did not receive the reference standard; thus, data were only available for 36 participants
Comparative	
Notes	There was no mention of funding source, but since this was a retrospective study, it is likely that the study received no explicit funding Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Unclear
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference stan-	Yes		

Uskul 2009 (*Continued*)

dard?			
Was there a pre-specified cut-off value?	Unclear		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Usuda 2013

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	158 participants, mean age = 68 (range = 37 to 83) years, 94 males/64 females, Japan **Histology of primary tumour** Adenocarcinoma: N = 114; squamous cell carcinoma: N = 36; large cell carcinoma: N = 3; small cell carcinoma: N = 3; adenosquamous carcinoma: N = 1; large cell neuroendocrine carcinoma: N = 1; carcinoid: N = 1; carcinosarcoma: N = 1 (2 participants had double lung cancers); comorbidities: not reported **Inclusion criteria**

	All participants with operable lung cancer or suspected operable lung cancer who agreed to DWI and PET-CT examinations were enrolled before operation in this study during the period from May 2009 to October 2010 **Exclusion criteria** Participants with metal or pacemakers in their body or tattoos on the skin (contraindication to MRI), bulky N2 lung cancers on PET or DWI N2 positive lymph nodes, or prior treatment for lung cancer **Previous tests** None listed **Clinical setting** Secondary/tertiary The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	The PET-CT scanning was performed with a dedicated PET camera (Siemens Biography Sensation 16) before surgery. All participants fasted for 6 hours before scanning. The FDG (185 MBq/3.7 MBq per kg) was administered intravenously. After a 60-minute uptake period, an emission scan was acquired for 3 minutes per bed position, and a whole-body scan was performed on each participant using several bed positions according to the height of each participant. 1 radiologist with 10 to 12 years of radioisotope scintigraphy and PET(-CT) experience, who was unaware of the participants' clinical data, and 1 pulmonologist with 28 years of experience evaluated the PET-CT data. A consensus was reached if there were any differences of opinion. After image reconstruction, a 2-dimensional round ROI was drawn on a slice after visual detection of the highest count on the fused CT image. For the lesions with negative or faintly positive PET findings, the ROI was drawn on the fusion image with the corresponding CT. From those ROI, the maximum standardised uptake value (SUVmax) was calculated as the FDG accumulation within primary lung cancers and lymph nodes **Covariates** Type of PET-CT scanner: Siemens Biography Sensation 16 FDG dose: 185 MBq / 3.7 MBq per kg Injection-to-scan time: 60 min Attenuation correction: not reported Cut-off values for test positivity (malignancy): A ROC curve was constructed according to the SUVmax using GraphPad Prism, and the cut-off values for a diagnosis of metastasis were determined. Lymph nodes with a SUVmax of the same or more than the cut-off value were defined as positive by means of PET-CT. Lymph nodes with a SUVmax less than the cut-off value or those that could not be detected on PET-CT were defined as negative by means of PET-CT. The cut-off value appears to be SUVmax $\geq$ or > 4.45
Target condition and reference standard(s)	Pathological staging (7 pneumonectomies, 1 bilobectomy, 119 lobectomies, 3 segmentectomies, and 30 partial resections). Pneumonectomies, bilobectomies, and lobectomies, but not segmentectomies and partial resections, were performed with systematic lymphadenectomies of the hilum and the mediastinal areas. It is assumed that smaller tumours did not have nodal metastases as no nodal sampling was performed
Flow and timing	Data were reported for 160 lung cancers in 158 participants
Comparative	

Notes	The study was supported partly by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sport, Science and Technology, Japan (21591828) Adverse events: not reported

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			

Usuda 2013 (*Continued*)

Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Yang 2008

Study characteristics	
Patient sampling	Prospective consecutive? patient series
Patient characteristics and setting	122 participants, median age = 69 (range = 32 to 84) years, 78 males/44 females, China **Histology of primary tumour** Adenocarcinoma: N = 63; squamous cell: N = 54; bronchoalveolar carcinoma: N = 3; large cell neuroendocrine cancer: N = 2; comorbidities: not reported **Inclusion criteria** Participants with histologically diagnosed or suspected non-small cell lung cancer referred for operation during the study period. The participants had stage I, II, and selected IIIA (i.e., the participants with T3 N1 stage and N2 stage with single stationed small N2 nodal metastasis or with only a few small N2 lymph node metastases after mediastinoscopy) **Exclusion criteria** Allergy to iodine contrast and hyperglycaemia over 9 mmol/L on the day when the FDG PET-CT scan was performed, participants who received prior chemotherapy or radiotherapy **Previous tests** Standard preoperative staging procedures, including physical examination, laboratory testing, and ultrasound of the neck and abdomen, chest radiography, pulmonary function, and bronchoscopy **Clinical setting** Cancer hospital The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	All participants fasted for at least 6 h before examination and had a normal blood glucose levels in peripheral blood. The participants then received an intravenous injection of 370 MBq (10 mCi) of FDG and rested for approximately 60 min before scanning. Scanning was performed with an integrated in-line PET/CT system (Discovery LS; GE Healthcare). Unenhanced CT was performed first followed by PET emission scan. 2 experienced nuclear medicine physicians were responsible for the interpretation of PET-CT images. Only 1 final decision was made by these 2 physicians **Covariates** Type of PET-CT scanner: integrated PET-CT scanner (Discovery LS; GE Healthcare) FDG dose: 370 MBq (10 mCi)

	Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): When an area of presumed lymph node showed FDG uptake that was focally prominent compared with surrounding tissues and not related to normal physiologic uptake, it was considered to be positive for malignancy. A site of increased FDG uptakes was defined as negative when it was related to the physiologic biodistribution of FDG
Target condition and reference standard(s)	100 participants underwent lobectomy, 12 participants underwent bilobectomy, 10 participants underwent pneumonectomy. During surgery, experienced thoracic surgeons dissected all visible and palpable lymph nodes in the surgical field that were accessible in the hilum and mediastinum, taking into consideration all results from the preoperative imaging examinations including the results of CT and PET/CT, irrespective of the size of the node. The specimens were stained by a standard haematoxylin-eosin staining and examined with optical light microscopy
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	This study was supported by Research Fund of Shandong Provincial Health Bureau, Shandong Province, China Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		

Yang 2008 (*Continued*)

Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Yang 2010

Study characteristics	
Patient sampling	Prospective consecutive (?) patient series
Patient characteristics and setting	31 participants, mean age = 59 (range = 38 to 84) years, 22 males/9 females, China **Histology of primary tumour** Adenocarcinoma: N = 13; squamous cell carcinoma: N = 11; adenosquamous carcinoma: N = 3; bronchoalveolar: N = 2; large cell neuroendocrine cancer: N = 2; comorbidities: not reported **Inclusion criteria** Participants with NSCLC **Exclusion criteria** Allergy to iodine contrast and hyperglycemia over 9 mmol/L on the day when the FDG PET-CT scan was performed, participants with other extra-thoracic metastasis or who had received prior chemotherapy or radiotherapy **Previous tests**

	Conventional lung cancer staging on the basis of clinical information and both FLT- and FDG PET-CT studies **Clinical setting** Cancer hospital The inclusion of only participants who received surgery narrows the range of patients who would receive PET-CT in practice, namely, patients (clinically) with suspected resectable non-small cell lung cancer, a proportion of whom would have N2 or N3 disease already on PET-CT
Index tests	All participants fasted for at least 6 h before examination, then received an intravenous injection of 300 to 400 MBq of FDG and rested for approximately 60 min before scanning. Scanning was performed with an integrated in-line PET-CT system (Discovery LS; GE Healthcare). Unenhanced CT was performed first, from the head to the thigh, with the following settings: 140 kV; 80 mA; tube rotation time, 0.5 s per rotation; a pitch of .75; and section thickness, 4.25 mm, which match the PET section thickness. A PET emission scan was performed that covered the identical transverse field of view immediately after CT. Acquisition time for PET was 4 min per table position. Participants were in normal shallow respiration during the image acquisition. PET data sets were reconstructed iteratively using CT data for attenuation correction, and coregistered images were displayed on a workstation (Xeleris; GE Healthcare). 2 experienced nuclear medicine physicians, unaware of surgical or pathological findings and any clinical information except for the participants with NSCLC, prospectively interpreted the PET-CT images. Tumour lesions were identified as areas of focally increased uptake exceeding that of the surrounding normal tissue. For primary tumours visualised on PET, a region of interest (ROI) was placed over the entire FDG-avid lesion on all transverse planes in which the tumour appeared and the SUVmax was calculated. For lesions not visible on the PET scan, an ROI was drawn on the scan corresponding to the area of abnormality on the CT image **Covariates** Type of PET-CT scanner: integrated PET-CT scanner (Discovery LS; GE Healthcare) FDG dose: 300 to 400 MBq Injection-to-scan time: 60 min Attenuation correction: yes Cut-off values for test positivity (malignancy): Lesions were considered positive if a definite localised area of higher FDG uptake than in the surrounding normal tissue was present, excluding physiologic uptake. There was no prespecified cut-off value
Target condition and reference standard(s)	Pathological results from lung cancer surgery performed within 2 weeks of PET-CT
Flow and timing	All participants were accounted for in the results. All participants received the reference standard. There were no uninterpretable results
Comparative	
Notes	Supported by the Research Fund of Shandong Provincial Health Bureau of China (grant 2009HZ088) and by the Research Fund of Shandong Cancer Hospital and Institute (no. 2009 to 11) Adverse events: not reported
Methodological quality	

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Yi 2008

Study characteristics	
Patient sampling	Prospective consecutive patient series
Patient characteristics and setting	165 participants, mean age = 61 (SD = 10, range = 34 to 82) years, 125 males/40 females, South Korea **Histology of primary tumour** Adenocarcinoma: N = 86; squamous cell: N = 59; NSCLC NOS: N = 9; pleomorphic carcinoma: N = 3; adenosquamous cell carcinoma: N = 3; sarcomatoid carcinoma N = 3; large cell neuroendocrine carcinoma: N = 2; comorbidities: not reported **Inclusion criteria** Consecutive participants with pathologically confirmed NSCLC who underwent both PET-CT (Discovery LS; GE Medical Systems, Milwaukee, WI, USA) and whole-body MR imaging (Achieva; Philips Medical Systems, Best, the Netherlands) at a single tertiary referral hospital from July 2005 to August 2006 and who did not have bronchioloalveolar cell carcinoma and who were not suspected of having metastatic disease at clinical examination (physical examination, laboratory findings at admission, and enhanced thoracic CT scans covering the thorax and upper abdomen) **Exclusion criteria** None listed, but 9 participants were excluded due to contraindications to MR scanning or inability to complete MR scanning (due to claustrophobia) **Previous tests** At least physical examination, laboratory findings at admission, and enhanced thoracic CT scans covering the thorax and upper abdomen **Clinical setting** Tertiary setting
Index tests	The glucose level in the peripheral blood was 150 mg/dL (8.33 mmol/L) or lower in all participants. Participants received an intravenous injection of 370 MBq (10 mCi) of FDG followed by rest for > 45 minutes before undergoing scanning. Scans were acquired by using a PET/CT device (Discovery LS; GE Healthcare, Milwaukee, WI, USA), which consisted of a PET scanner (Advance NXi; GE Healthcare) and an eight-section CT scanner (Light-Speed Plus; GE Healthcare). Immediately after unenhanced CT was performed, emission PET was performed in the identical transverse field of view. Integrated PET/CT images were evaluated jointly by a chest radiologist (with 18 years of CT interpretation experience) and a nuclear medicine physician (with 4 years of PET-CT interpretation experience) with consideration of the diagnostic criteria listed below (under **Covariates**) for a positive finding. No details were included about attenuation correction **Covariates** Type of PET-CT scanner: PET/CT device (Discovery LS; GE Healthcare, Milwaukee, WI, USA)

	FDG dose: 370 MBq (10 mCi) Injection-to-scan time: > 45 min Attenuation correction: not reported Cut-off values for test positivity (malignancy): All lymph nodes in the thorax with abnormal FDG uptake (greater than mediastinal blood pool uptake) were considered to be metastatic, unless they showed high attenuation (70 HU) or benign calcification (central nodular, laminated, popcorn-like, or diffuse) on unenhanced CT images. An abnormal focal FDG uptake that accompanied a corresponding anatomic alteration was considered to be indicative of metastasis. Subcentimeter lung nodules, regardless of FDG uptake (usually less than mediastinal blood pool uptake), were considered malignant when they were greater than 10 in number. Any nodule more than 10 mm in diameter with FDG uptake more than that of the mediastinal blood pool was considered malignant. Both image readers were unaware of whole-body MR imaging findings and of clinical and pathologic evaluation results
Target condition and reference standard(s)	N stage was determined in 150/165 participants, with thoracotomy results in 118, mediastinoscopy results in 26, and supraclavicular nodal biopsy results in 6
Flow and timing	N-stage was available for 150/165 participants. It was unclear why N-stage is not available for the remaining 15 participants
Comparative	
Notes	This study was supported by Samsung Medical Center Clinical Research Development Program grant #CRDP CRS 106-41-2. The authors stated that they have no financial relationship to disclose Adverse events: not reported

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Was there a pre-specified cut-off value?	No		
Was a positive result defined?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
			Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

CT = computed tomography.
DW = diffusion weighted.
DWI = diffusion weighted imaging.
EBUS-TBNA = endobronchial ultrasound-guided transbronchial needle aspiration.
ECAT ART = a PET scanner equipped with a partial ring of bismuth germanate detectors.
ECOG = Eastern Cooperative Oncology Group.
EUS = endoscopic transoesophageal ultrasound.
EUS-FNA = endoscopic ultrasound-guided fine needle aspiration.

FDG = (^{18} F)-2-fluoro-deoxy-D-glucose.
FLT-PET/CT = 18-fluorothymidine-PET/CT.
FN = false negative.
FP = false positive.
h = hours.
IV = intravenous.

LN = lymph node.
M1 = metastatic.
max = maximum.
mins = minutes.
MLN = mediastinal lymph node.
MR = magnetic resonance.
MRI = magnetic resonance imaging.
NOS = not otherwise specified.
N = number.
NSCLC = non-small cell lung cancer.
OS-EM = ordered-subset expectation-maximisation.
OS-EM = ordered-subset expectation-maximisation.
PET = computed tomography.
PET-CT = positron emission tomography-computed tomography.
RI = retention index.
ROI = region of interest.
SCLC = small cell lung cancer.
SD = standard deviation.
STIR = short inversion time inversion-recovery.
SUV = standardised uptake value.
Syngo = medical imaging software
TB = tuberculosis.
TBNA = transbronchial needle aspiration.
TN = true negative.
TNM = TNM classification of malignant tumours.
TP = true positive.
VATS = video-assisted thoracoscopic surgery.
Please visit the following websites for information about the abbreviations used for units: www1.bipm.org/en/measurement-units/ (International Bureau of Weights and Measures); physics.nist.gov/cuu/Units/outside.html (The NIST Reference on Constants, Units, and Uncertainty).

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Agraval 2012	2-by-2 table could not be extracted for patient-level N-staging data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Ahmed 2010	Conference abstract. 2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 15 November 2012
Akpinar 2013	Index test scanning employed an ECAT Accel scanner (Siemens, Erlangen, Germany), which is a dedicated PET scanner, that is, not an integrated PET/CT scanner
Al-Ibraheem 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Al-Sarraf 2008	The unit of analysis was lymph node. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012

(*Continued*)

Allen-Auerbach 2006	29/142 included participants had NSCLC. The remainder had other types of cancer. 2-by-2 table could not be extracted for this subgroup. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012. Author responded that study was not set up for assessing the accuracy of PET-CT for staging of NSCLC. The aim of the study was solely to determine the incidence of missed pulmonary micronodules on PET/CT studies acquired during shallow breathing. Thus, the data relevant for this review were not part of this study
An 2008	The unit of analysis was lymph node. 2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Antoch 2003	N2 status of the participants could not be ascertained as participants were classified according to overall stage with no individual breakdown of T-, N-, and M status. Emailed for per-patient-based-data for N0 and N1 v N2 and N3 on 1 November 2012
Aquino 2003	FDG-PET and CT performed on different scanners within a 1-month time interval and subsequently coregistered
BalciI 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Beyer 2010	Reference standard consisted of a combination of FDG uptake in lymph nodes, findings from either histological examinations , FDG-PET/CT, or both, and clinical follow-up examinations for a minimum of 2 months
Bhatt 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Booth 2009	Conference abstract. 2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Boulougouri 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Bryant 2006b	2-by-2 table could not be extracted. Emailed for per-patient-based data for N0-1 v N2 and N3 on 1 November 2012
Carrillo 2012	2-by-2 table could not be extracted for patient level N0-1 v N2 and N3 data. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013. Correspondance with the author clarified that N1 and N2 disease were treated as N-positive as a whole and not distinguished between
Cerfolio 2004	13 participants with pathological N0 disease were over-staged by FDG-PET/CT. Not enough information was provided to ascertain which N-stage these participants were assigned to by FDG-PET/CT. It was therefore not possible to know how many of these participants were true negatives (if staged by FDG-PET/CT as N1 disease) and how many of these participants were false positives (if staged as N2 or N3 disease by FDG-PET/CT). Similarly, 2 participants with pathological N3 disease were under-staged by FDG-PET/CT. Not enough information was provided to ascertain which N-stage these participants were staged as by FDG-PET/CT, and it is therefore impossible to know whether these participants should be classed as true positives (if staged as N2 disease by FDG-PET/CT) or false negatives (if staged as N0 or N1 disease by FDG-PET/CT). Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012

(*Continued*)

Cerfolio 2006	Included only participants who were clinically N0 and N1 and excluded all participants who were clinically N2, N3, and M1 even if those test results later turned out to be false positive, which meant no false positive results would be possible in the included data set
Cerfolio 2007	Unit of analysis = node. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Cerfolio 2008	Data can only be verified for 14 to 26/166 participants. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013
Cetinkaya 2011	Unclear if all participants received PET-CT. 2-by-2 could not be extracted. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013
Ceylan 2012	2-by-2 could not be extracted. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013
Chiba 2010	No pathological reference standard
Colville 2013	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Cömert 2012	2-by-2 could not be extracted. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013
Delgado-Bolton 2010	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Duan 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Faber 2011	Conference abstract. 2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Flechsig 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Gregory 2012	61/168 participants received the reference standard
Gómez-Caro 2012	Included only participants who were clinical stage 1 and excluded all participants who were clinical stage IIA and above, which meant no false positive results would be possible in the included data set
Günlüoğ lu 2010	2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Günlüoğ lu 2011b	Published in Turkish. Unsure if it is PET or PET-CT. 2-by-2 table could not be extracted. Emailed author for clarification and per-patient-based data for N0 and N1 v N2 and N3 on 18 March 2013

(*Continued*)

Halpern 2005	2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Hong 2010	Conference abstract. 2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Hu 2008	2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Huang 2012	Included only PET-CT-negative participants, so no TP or FP possible
Kasai 2010	2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Kelly 2006	18/49 participants did not receive reference standard
Kim 2011	Included only participants who were staged N0 and N1 on both CT and PET-CT, and excluded all participants who were staged N2 and N3 or M1 on CT or PET-CT, which meant no false positive results would be possible in the included data set
Kim 2012	2-by-2 could not be extracted. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013
Kim 2012a	9/69 participants received the reference standard
Kim 2012b	Included only PET/CT-negative participants, so no possibility of TPs or FPs
Kommata 2011	337/401 participants did not receive reference standard
Krueger 2006	Conference abstract. 2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 5 November 2012
Lapinska 2011	Conference abstract. 2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Lardinois 2003	2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Lasnon 2012	25/71 participants received reference standard
Lebioda 2013	Only PET-CT N2-negative nodes were investigated, thus, no TPs or FPs possible
Lee 2004	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013, but received delivery failure report on email
Lee 2008	Unit of analysis = node. 2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012

(*Continued*)

Lee 2009b	During the study period 63 newly diagnosed NSCLC participants underwent both CT and integrated PET-CT for clinical staging. Of those, 43 participants underwent surgical procedures (including mediastinoscopy) , which allowed pathologic evaluation of the mediastinal nodes, and only these 43 participants were included in the study, of whom data can only be ascertained for 37
Li 2009	2-by-2 table could not be extracted for patient-level N-staging data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Li 2011	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Li 2012b	2-by-2 could not be extracted. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013
Li 2012c	Only PET-CT positive nodes were investigated, thus, no TNs or FNs possible
Lin 2012	2-by-2 could not be extracted. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013
Liu 2009	Unit of analysis = node. 2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Low 2006	17/41 participants received the gold standard
Ma 2011	2-by-2 table could not be extracted for patient-level N-staging data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Maeda 2009	Included only participants who were clinical stage IA, which meant no false positive results would be possible in the included data set
Mariam 2009	Conference abstract. 2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 5 November 2012
Meduoye 2009	Not integrated FDG-PET/CT, but FDG-PET and CT performed separately
Mendez 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Mi 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Moodie 2009	No reference standard
Moreno Garcia 2009	Conference abstract. 2-by-2 table could not be extracted. Could not find author contact details to request per-patient-based data for N0 and N1 v N2 and N3
Morikawa 2011	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013

(*Continued*)

Nakajima 2011	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Nakamura 2008	Reference test consisted of clinical course or pathology findings in biopsy specimens from informative sites for non-surgical cases (N = 30/50)
Nomori 2008	2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2-3 on 1 November 2012
Ohno 2007	N0 disease classified as test negative and N1, N2, and N3 disease classified as test positives. Not enough information was provided to permit reclassification of N1 disease to test negatives. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Ozkan 2010	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Pauls 2012	None of the participants received the reference standard
Pfannenberg 2007	Population consists of 50 participants referred for PET/CT for primary tumour staging before surgery (N = 4), neoadjuvant radiochemotherapy (N = 24), definitive radiotherapy (N = 8), or palliative radiotherapy (N = 5). The remaining 9 participants were referred for restaging during follow-up because of suspected local or distant recurrence Reference standard consisted of surgical staging for 35 participants and follow up for 15 participants
Pozo-Rodriguez 2005	Not integrated FDG-PET/CT, but FDG-PET and contrast-enhanced CT performed separately on different scanners and not coregistered
Prévost 2009	2-by-2 table could not be extracted for the 44 participants with pathological staging. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Quaia 2008	N0 disease classified as test negative and N1, N2, and N3 disease classified as test positives. Not enough information was provided to permit reclassification of N1 disease to test negatives. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Salman 2009	Conference abstract. 2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Schiavariello 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Schreyögg 2010	77/172 participants received reference standard
Schwenzer 2012	No reference standard
Shim 2005	Retrospective study: unit of analysis = node Prospective study: unit of analysis = node. Overall staging results provided for stages I, II, III, and IV, and 2-by-2 table for N-staging with participant as unit of analysis could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012

(*Continued*)

Sit 2010	2-by-2 table could not be extracted for patient-level N-staging data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Sivrikoz 2010	Only PET-CT-negative nodes were investigated, thus, no TPs or FPs possible
Sivrikoz 2012	2-by-2 could not be extracted for all participants. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013
Steinert 2010	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Sánchez Sánchez 2011	14/34 participants did not received pathological confirmation of PET/CT results
Tamura 2012	2-by-2 could not be extracted. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013
Tsutani 2012	Included only PET/CT-negative participants, i.e., no chance of TPs or FPs
Vaz 2012	2-by-2 could not be extracted. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25/2/13
Ventura 2010	Unit of analysis = node. 2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012, but received delivery failure report. Re-sent on 5 November 2012 with the same result. We could not find any other contact details for the corresponding author
Wang 2012	2-by-2 table could not be extracted for patient-level N-staging data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Wiese 2012	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013 using the "contact us" webform on www.leading-medicine-guide.ch/en/spitzenmediziner/kontakt/id/2492
Wu 2010	2-by-2 table could not be extracted for patient-level N-staging data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 13 June 2013
Yi 2007	The participants appear to be retrospectively recruited from the same institution as Shin 2008 during the period of July 2003 until June 2005. Shin 2008 recruited (also retrospectively) their participants during the period September 2003 until July 2006, which means the samples were largely overlapping. We retained Shin Shin 2008 because this study has the larger sample of the 2 studies
Yi 2011	Abstract. Unit of analysis = node. 2-by-2 table could not be extracted for patient-level N-staging data. Emailed via Kyung Soo Lee for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Yi 2013	2-by-2 table could not be extracted for patient-level data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 10 July 2013

(*Continued*)

Yu 2007	Paper primarily about diagnosis, not staging, with 34 out of 104 participants having pathological diagnosis and 70 out of 104 participants having clinical diagnosis. 18 of the 104 participants had benign lesions, and 86 participants had lung cancer. 12 of 25 lung cancer participants with no metastasis on FDG-PET/CT received surgery
Zhang 2006	2-by-2 table could not be extracted. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012, but received delivery failure report. Re-sent on 7 November 2012 with the same result. We could not find any other contact details for the corresponding author
Zhang 2012	2-by-2 table could not be extracted for patient-level N-staging data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Zsiray 2009	Hungarian. Describes their experience and there appears to be no consistent reference standard, and no possibility of extracting a 2-by-2 with clarity that it is N0 and N1 v N2 and N3. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 1 November 2012
Zsiray 2011	Not pathological reference standard
Zsiray 2012	2-by-2 table could not be extracted for patient-level N-staging data. Emailed for per-patient-based data for N0 and N1 v N2 and N3 on 17 June 2013
Özta 2012	2-by-2 could not be extracted. Emailed author for per-patient-based data for N0 and N1 v N2 and N3 on 25 February 2013

FDG = (^{18} F)-2-fluoro-deoxy-D-glucose.

FDG-PET/CT = (^{18} F)-2-fluoro-deoxy-D-glucose positron emission tomography-computed tomography.
FPs = false positives.
NSCLC = non-small cell lung cancer.
TPs = true positives.
v = versus.

DATA

Presented below are all the data for all of the tests entered into the review.

Tests. Data tables by test

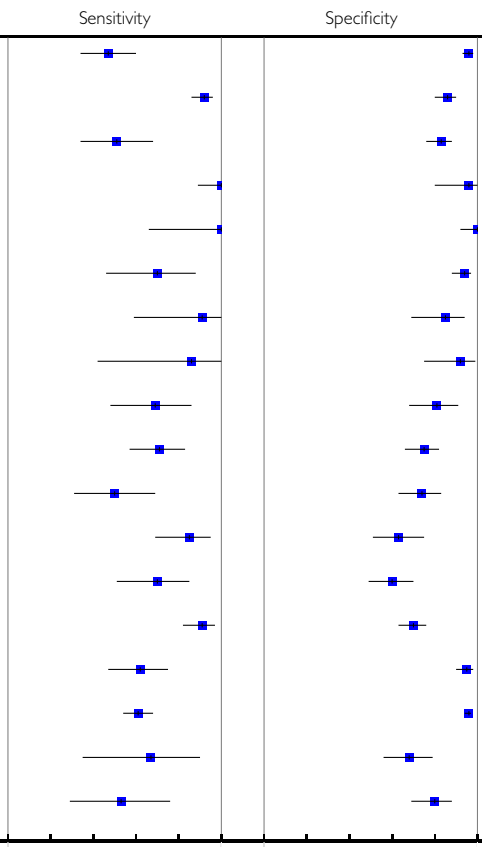
Test	No. of studies	No. of participants
1 PET/CT	45	6095

Test 1. PET/CT.

Review: PET-CT for assessing mediastinal lymph node involvement in patients with suspected resectable non-small cell lung cancer

Test: 1 PET/CT

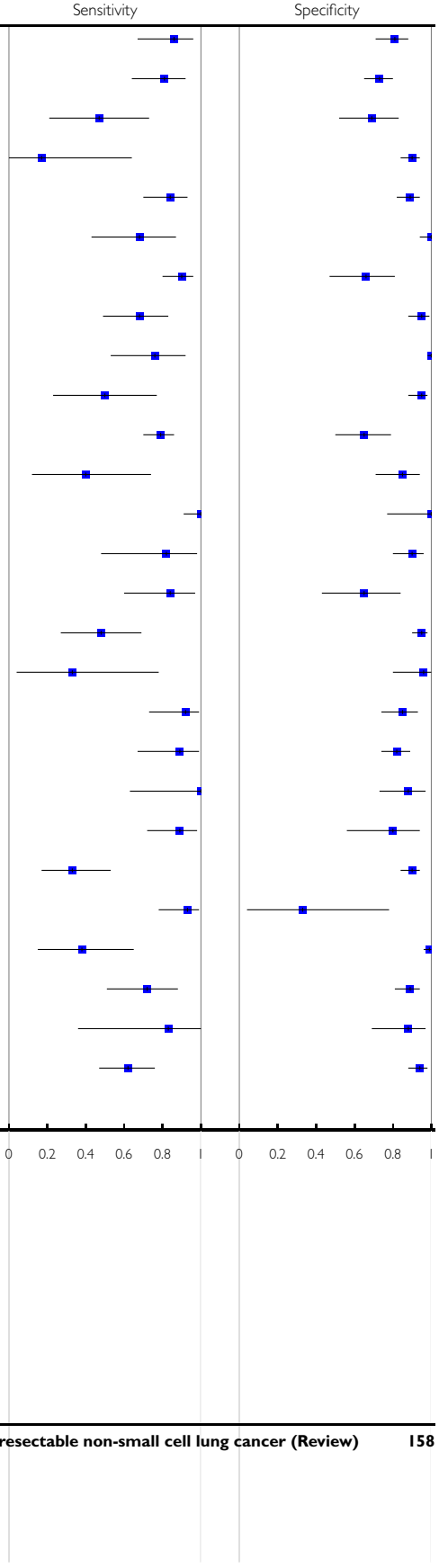
Study	TP	FP	FN	TN	Sensitivity	Specificity
Bille 2013	28	13	32	280	0.47 [0.34, 0.60]	0.96 [0.93, 0.98]
Bryant 2006a	131	33	12	195	0.92 [0.86, 0.96]	0.86 [0.80, 0.90]
Carnochan 2009	19	27	18	130	0.51 [0.34, 0.68]	0.83 [0.76, 0.88]
Chen 2010	31	1	0	24	1.00 [0.89, 1.00]	0.96 [0.80, 1.00]
Czepczynski 2011	9	0	0	42	1.00 [0.66, 1.00]	1.00 [0.92, 1.00]
Darling 2011	14	8	6	121	0.70 [0.46, 0.88]	0.94 [0.88, 0.97]
De Wever 2007	10	6	1	33	0.91 [0.59, 1.00]	0.85 [0.69, 0.94]
El-Hariri 2012	6	2	1	24	0.86 [0.42, 1.00]	0.92 [0.75, 0.99]
Fischer 2011	18	10	8	43	0.69 [0.48, 0.86]	0.81 [0.68, 0.91]
Gunluoglu 2011	35	30	14	89	0.71 [0.57, 0.83]	0.75 [0.66, 0.82]
Harders 2012	15	22	15	62	0.50 [0.31, 0.69]	0.74 [0.63, 0.83]
Hu 2011	29	25	5	43	0.85 [0.69, 0.95]	0.63 [0.51, 0.75]
Hwangbo 2009	21	35	9	52	0.70 [0.51, 0.85]	0.60 [0.49, 0.70]
Iskender 2012	63	65	6	152	0.91 [0.82, 0.97]	0.70 [0.63, 0.76]
Jeon 2010	32	8	20	150	0.62 [0.47, 0.75]	0.95 [0.90, 0.98]
Kim 2007	110	21	70	473	0.61 [0.54, 0.68]	0.96 [0.94, 0.97]
Koksal 2013	8	22	4	47	0.67 [0.35, 0.90]	0.68 [0.56, 0.79]
Kuo 2012	10	17	9	66	0.53 [0.29, 0.76]	0.80 [0.69, 0.88]



(*Continued* . . .)

(. . . *Continued*)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Lee 2007	24	19	4	79	0.86 [0.67, 0.96]	0.81 [0.71, 0.88]
Lee 2009a	29	40	7	106	0.81 [0.64, 0.92]	0.73 [0.65, 0.80]
Lee 2011	7	12	8	27	0.47 [0.21, 0.73]	0.69 [0.52, 0.83]
Lee 2012	1	16	5	138	0.17 [0.00, 0.64]	0.90 [0.84, 0.94]
Li 2010	41	12	8	97	0.84 [0.70, 0.93]	0.89 [0.82, 0.94]
Li 2012a	13	0	6	61	0.68 [0.43, 0.87]	1.00 [0.94, 1.00]
Morikawa 2009	55	11	6	21	0.90 [0.80, 0.96]	0.66 [0.47, 0.81]
Ohnishi 2011	21	4	10	75	0.68 [0.49, 0.83]	0.95 [0.88, 0.99]
Ohno 2011	16	0	5	229	0.76 [0.53, 0.92]	1.00 [0.98, 1.00]
Ose 2012	7	5	7	93	0.50 [0.23, 0.77]	0.95 [0.88, 0.98]
Ozkan 2011	84	16	23	30	0.79 [0.70, 0.86]	0.65 [0.50, 0.79]
Perigaud 2009	4	6	6	35	0.40 [0.12, 0.74]	0.85 [0.71, 0.94]
Plathow 2008	38	0	0	14	1.00 [0.91, 1.00]	1.00 [0.77, 1.00]
Sanli 2009	9	7	2	60	0.82 [0.48, 0.98]	0.90 [0.80, 0.96]
Saydam 2012	16	8	3	15	0.84 [0.60, 0.97]	0.65 [0.43, 0.84]
Shin 2008	11	8	12	153	0.48 [0.27, 0.69]	0.95 [0.90, 0.98]
Sommer 2012	2	1	4	24	0.33 [0.04, 0.78]	0.96 [0.80, 1.00]
Subedi 2009	22	10	2	57	0.92 [0.73, 0.99]	0.85 [0.74, 0.93]
Tasci 2010	17	19	2	89	0.89 [0.67, 0.99]	0.82 [0.74, 0.89]
Toba 2010	8	4	0	30	1.00 [0.63, 1.00]	0.88 [0.73, 0.97]
Tournoy 2007	25	4	3	16	0.89 [0.72, 0.98]	0.80 [0.56, 0.94]
Uruga 2011	10	15	20	137	0.33 [0.17, 0.53]	0.90 [0.84, 0.94]
Uskul 2009	28	4	2	2	0.93 [0.78, 0.99]	0.33 [0.04, 0.78]
Usuda 2013	6	1	10	143	0.38 [0.15, 0.65]	0.99 [0.96, 1.00]
Yang 2008	18	11	7	86	0.72 [0.51, 0.88]	0.89 [0.81, 0.94]
Yang 2010	5	3	1	22	0.83 [0.36, 1.00]	0.88 [0.69, 0.97]
Yi 2008	28	6	17	99	0.62 [0.47, 0.76]	0.94 [0.88, 0.98]



ADDITIONAL TABLES

Table 1. Study details

Study ID	First author, year of publication
Patient sampling	Prospective/retrospective, case-control/consecutive/random patient series
Patient characteristics and setting	Inclusion and exclusion criteria, previous tests for lung cancer (diagnosis and staging), clinical setting, sample size, age, sex, comorbidities, country, histology of primary tumour
Index test	Details of the index test used including the type of PET-CT scanner, FDG dose, injection-to-scan time, attenuation correction, and the cut-off values for test positivity (malignancy)
Reference standard(s)	The reference standard(s) used
Flow and timing	All participants were accounted for in results, missing/uninterpretable test results, reasons for withdrawal, adverse events caused by test, the time interval, and any interventions between index test(s) and reference standard
Notes	Source of funding, anything else of relevance

FDG = (^{18} F)-2-fluoro-deoxy-D-glucose.
PET-CT = positron emission tomography and computed tomography.

Table 2. Risk of bias and applicability items and criteria for their assessment

Item	Description
Domain 1: Patient selection	-
A. Risk of bias	-
Patient sampling	The study design will be listed here
Was a consecutive or random sample of participants enrolled?	'Yes' if a consecutive or random sample of participants were enrolled 'No' if a consecutive or random sample of participants were not enrolled 'Unclear' if the study does not describe the method of participants enrolment
Was a case-control design avoided?	'Yes' if the study has not used a case-control design 'No' if the study has used a case-control design 'Unclear' if the study does not report enough information to ascertain whether a case-control design was used

Table 2. Risk of bias and applicability items and criteria for their assessment (*Continued*)

Did the study avoid inappropriate exclusions?	'Yes' if the characteristics of the participants are well described and probably typical of a secondary healthcare setting 'No' if the sample is unrepresentative of people with potentially resectable lung cancer in general 'Unclear' if the source or characteristics of participants is not adequately described
Could the selection of participants have introduced bias?	A judgement of low, high, or unclear risk of bias will be made based on a balanced assessment of the responses to the above signalling questions
B. Concerns about applicability	-
Patient characteristics and setting	The information detailed under 'Patient characteristics and setting' in the 'Characteristics of included studies' table will be listed here
Are there concerns that the included participants and setting do not match the review question?	A judgement of low, high, or unclear concerns about applicability will be made based on a balanced assessment of the response to the third signalling question above and on how closely the sample matches the target population of interest
Domain 2: Index test	-
Index test	The information detailed under 'Index test' in the 'Characteristics of included studies' table will be listed here
Were the index test results interpreted without knowledge of the results of the reference standard?	'Yes' if the report stated that the person undertaking the index test did not know the results of the reference tests or if the 2 tests were carried out in different places 'No' if the report stated that the same person performed both tests or that the results of the index tests were known to the person undertaking the reference tests 'Unclear' if insufficient information provided
Did the study provide a clear definition of what was considered to be a positive result?	'Yes' if the definition of a positive result is clearly stated (e.g., SUV) 'No' if no definition of what was considered a positive result is stated or the definition of a positive result varied between the participants 'Unclear' if not enough information is given to permit judgement
If a threshold was used, was it prespecified?	'Yes' if prespecified 'No' if the authors selected the optimal cut-off value based on the results of the study 'Unclear' if there is a range of cut-off values and there is doubt which cut-off has been used, or if there is no mention at all of a cut-off value

Table 2. Risk of bias and applicability items and criteria for their assessment (*Continued*)

Could the conduct or interpretation of the index test have introduced bias?	A judgement of low, high, or unclear risk of bias will be made based on a balanced assessment of the responses to the above signalling questions
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	A judgement of low, high, or unclear concerns about applicability will be made based on a balanced assessment of the information detailed under 'Index test' with particular reference to the definition of test positivity/malignancy
Domain 3: Reference standard	-
Target condition and reference standard(s)	The information detailed under 'Reference standard(s)' in the 'Characteristics of included studies' will be listed here
Is the reference standards likely to correctly classify the target condition?	'Yes' if reference standard is sampling of mediastinal nodes with pathological diagnosis 'No' if there is no sampling of mediastinal nodes with pathological diagnosis 'Unclear' if insufficient information is provided
Were the reference standard results interpreted without knowledge of the results of the index tests?	'Yes' if the report stated that the person undertaking the index test did not know the results of the reference tests, or if the 2 tests were carried out in different places 'No' if the report stated that the same person performed both tests, or that the results of the index tests were known to the person undertaking the reference tests 'Unclear' if insufficient information provided
Could the reference standard, its conduct, or its interpretation have introduced bias?	A judgement of low, high, or unclear risk of bias will be made based on a balanced assessment of the responses to the above signalling questions
Are there concerns that the target condition as defined by the reference standard does not match the question?	The answer to this question will always be 'low' because the target condition that the reference standard defines will always be the target condition of the review, i.e., pathologically confined non-small cell lung cancer. Otherwise, the study will not be included
Domain 3: Flow and timing	-
Flow and timing	The information detailed under 'Flow and timing' in the 'Characteristics of included studies' will be listed here
Was there an appropriate interval between index test and reference standard?	'Yes' if the time period between PET-CT and the reference standard is $\leq$ 8 weeks 'No' if the time period between PET-CT and the reference standard is > 8 weeks 'Unclear' if insufficient information is provided

Table 2. Risk of bias and applicability items and criteria for their assessment (*Continued*)

Did all participants receive the same reference standard?	'Yes' if the same reference test was used regardless of the index test results 'No' if different reference tests are used depending on the results of the index test 'Unclear' if insufficient information is provided If any participants received a different reference test, what were the reasons stated for this, and how many participants were involved?
Were all participants included in the analysis?	'Yes' if there are no participants excluded from the analysis, or if exclusions are adequately described 'No' if there are participants excluded from the analysis and there is no explanation given 'Unclear' if not enough information is given to assess whether any participants were excluded from the analysis Report how many participants were excluded from the analysis for reasons other than uninterpretable results Report how many results were uninterpretable (of the total)
Could the patient flow have introduced bias?	A judgement of low, high, or unclear risk of bias will be made based on a balanced assessment of the responses to the above signalling questions and of the information listed under 'Flow and timing'
Was the study free of commercial funding?	'Yes' if the funding source is clearly stated and is not commercial 'No' if the funding source is clearly stated and is commercial 'Unclear' if not enough information is given to assess whether the funding source is commercial

PET-CT = positron emission tomography and computed tomography.
SUV = standardised uptake value.

Table 3. Sensitivity analysis of PET-CT accuracy for assessing mediastinal lymph node involvement

Covariate	N	Sen	(95% CI)	Spe	(95% CI)
Activity > background	18	77.4	(65.3 to 86.1)	90.1	(85.3 to 93.5)
Selection bias (low RoB)	6	67.2	(55.7 to 76.9)	89.7	(81.9 to 94.3)
Participants and setting (no concerns with applicability)	5	86.0	(55.2 to 96.8)	87.7	(77.7 to 93.5)
Index test bias (low RoB)	11	68.2	(59.1 to 76.1)	92.0	(87.4 to 95.0)

Table 3. Sensitivity analysis of PET-CT accuracy for assessing mediastinal lymph node involvement (*Continued*)

Index text (no concerns with applicability)	18	77.4	(65.3 to 86.1)	90.1	(85.3 to 93.5)
Reference standard (low RoB)	17	73.6	(63.0 to 82.0)	89.6	(84.6 to 93.1)
Reference standard bias (no concerns with applicability)	18	77.4	(65.3 to 86.1)	90.1	(85.3 to 93.5)
Flow and timing (low RoB)	10	77.0	(64.1 to 86.3)	90.5	(83.2 to 94.8)
Clear definition of test positivity (yes)	16	68.5	(60.7 to 75.4)	89.2	(84.1 to 92.7)
Non-commercial funding (yes)	6	62.0	(53.5 to 69.8)	92.8	(85.8 to 96.5)
SUVmax $\geq$ 2.5	12	81.3	(70.2 to 88.9)	79.4	(70.0 to 86.5)
Selection bias (low RoB)	6	81.4	(62.3 to 92.1)	81.8	(72.1 to 88.7)
Participants and setting (no concerns with applicability)	4	88.2	(79.7 to 93.5)	76.2	(63.9 to 85.2)
Index test bias (low RoB)	7	82.7	(66.6 to 92.0)	75.8	(62.2 to 85.7)
Index text (no concerns with applicability)	12	81.3	(70.2 to 88.9)	79.4	(70.0 to 86.5)
Reference standard bias (low RoB)	12	81.3	(70.2 to 88.9)	79.4	(70.0 to 86.5)
Reference Standard (no concerns with applicability)	11	79.6	(67.8 to 87.9)	81.1	(72.7 to 87.4)
Flow and timing (low RoB)	6	74.4	(55.6 to 87.0)	75.6	(63.6 to 84.6)
Clear definition of test positivity (yes)	12	81.3	(70.2 to 88.9)	79.4	(70.0 to 86.5)

Table 3. Sensitivity analysis of PET-CT accuracy for assessing mediastinal lymph node involvement (*Continued*)

Non-commercial funding (yes)	5	74.2	(54.6 to 87.3)	76.6	(61.4 to 87.1)

CI = confidence interval
N = number of participants
RoB = risk of bias
Sen = sensitivity
Spe = specificity

APPENDICES

Appendix 1. Glossary

Bifurcation: Where an airway or vessel divides into two.
Contiguous: Adjoining and touching, e.g., contiguous slices on CT scan means that they are adjoining so no space is missed in between.
Contralateral: The opposite side to the reference lesion, e.g., contralateral lymph nodes are those on the opposite side of the mediastinum to the main tumour.
Coronal: Parallel to the long axis of the body in the medial to later plane.
Hilum: The area where the blood vessels, airways, and nerves enter and leave the mediastinum.
Interstitial: The sponge-like substance of the lung.
Ipsilateral: The same side as the reference lesion.
Locoregional: Close to or in the immediate vicinity of the reference (a tumour or lesion).
Mediastinoscopy: A procedure where a rigid tube is inserted behind the top of the sternum and the structures in the mediastinum area are visualised and can be sampled.
Mediastinum: The mass of tissues and organs separating the two pleural sacs, between the sternum in front and the vertebral column behind, containing the heart and its large vessels, trachea, oesophagus, thymus, lymph nodes, and other structures and tissues; it is divided into superior and inferior regions, the latter subdivided into anterior, middle, and posterior parts.
Orbitomeatal: Line running from the external auditory meatus of the skull to the lower border of the orbit.
Paratracheal: Adjacent to the trachea.
Parenchyma: The sponge-like substance of the lung.
Peri-oesophageal: Around or adjacent to the oesophagus.
Sagittal: Parallel to the long axis of the body in the anterior to posterior plane.
Subcarinal: Immediately beneath the carina, the area where the trachea divides into the right and left main bronchi.
TNM system: A WHO recognised method for staging cancers according to the size and extent of the primary tumour (T; Tx: tumour cannot be evaluated; Tis: carcinoma in situ; T0: no signs of tumour; T1, T2, T3, T4: size and/or extension of the primary tumour), the degree of spread to regional lymph nodes (N; Nx: lymph nodes cannot be evaluated; N0: tumour cells absent from regional lymph nodes; N1: regional lymph node metastasis present; (at some sites: tumour spread to closest or small number of regional lymph nodes); N2: tumour spread to an extent between N1 and N3 (N2 is not used at all sites); N3: tumour spread to more distant or numerous regional lymph nodes (N3 is not used at all sites)) and presence of distant metastasis (M; M0: No distant metastasis; M1: metastasis to distant organs (beyond regional lymph nodes)).
Transcarinal: Through the carina, e.g., transcarinal needle aspiration is where a needle is inserted through the carina.

Appendix 2. Search strategies

MEDLINE/PreMEDLINE via OvidSP

1	exp Lung Neoplasms/
2	((lung or lungs or pulmonary) adj3 (neoplasm$ or cancer$ or carcinoma$ or adenocarcinoma$ or angiosarcoma$ or chrondosarcoma$ or sarcoma$ or teratoma$ or lymphoma$ or blastoma$ or microcytic$ or tumour$ or tumor$)).ti,ab
3	NSCLC.ti,ab.
4	or/1-3
5	Tomography/
6	Tomography, Emission-Computed/
7	Positron-Emission Tomography/
8	Tomography, Spiral Computed/
9	Fludeoxyglucose f 18/
10	FDG or Fludeoxyglucose or fluorodeoxyglucose or depreotide).tw
11	((positron or photon or scintillation) adj3 (emission or tomograph$)).tw
12	(CGC or PET or SPECT or NEOTECT or NEOSPECT or NEOTEC).tw.
13	or/5-12
14	4 and 13

EMBASE via OvidSP

#1	MeSH descriptor: [Lung Neoplasms] explode all trees
#2	((lung or lungs or pulmonary) near/3 (neoplasm* or cancer or carcinoma* or adenocarcinoma* or angiosarcoma* or chrondosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or tumour* or tumor*))
#3	NSCLC
#4	#1 or #2 or #3
#5	MeSH descriptor: [Tomography] this term only
#6	MeSH descriptor: [Tomography, Emission-Computed] this term only

(*Continued*)

#7	MeSH descriptor: [Positron-Emission Tomography] this term only
#8	MeSH descriptor: [Tomography, Spiral Computed] this term only
#9	MeSH descriptor: [Fluorodeoxyglucose F18] this term only
#10	(FDG or Fludeoxyglucose or fluorodeoxyglucose or depreotide)
#11	((positron or photon or scintillation) near/3 (emission or tomograph*))
#12	(CGC or PET or SPECT or NEOTECT or NEOSPECT or NEOTEC)
#13	#5 or #6 or #7 or #8 or #9 or #10 or #11 or #12
#14	#4 and #13

The Cochrane Library

#1	MeSH descriptor: [Lung Neoplasms] explode all trees
#2	((lung or lungs or pulmonary) near/3 (neoplasm* or cancer or carcinoma* or adenocarcinoma* or angiosarcoma* or chrondosarcoma* or sarcoma* or teratoma* or lymphoma* or blastoma* or microcytic* or tumour* or tumor*))
#3	NSCLC
#4	#1 or #2 or #3
#5	MeSH descriptor: [Tomography] this term only
#6	MeSH descriptor: [Tomography, Emission-Computed] this term only
#7	MeSH descriptor: [Positron-Emission Tomography] this term only
#8	MeSH descriptor: [Tomography, Spiral Computed] this term only
#9	MeSH descriptor: [Fluorodeoxyglucose F18] this term only
#10	(FDG or Fludeoxyglucose or fluorodeoxyglucose or depreotide)
#11	((positron or photon or scintillation) near/3 (emission or tomograph*))
#12	(CGC or PET or SPECT or NEOTECT or NEOSPECT or NEOTEC)
#13	#5 or #6 or #7 or #8 or #9 or #10 or #11 or #12

(*Continued*)

#14	#4 and #13

ProQuest Dissertation & Theses
(lung NEAR/3 (neoplasm OR cancer OR carcinoma OR adenocarcinoma OR angiosarcoma OR chrondosarcoma OR sarcoma OR teratoma OR lymphoma OR blastema OR microcytic OR tumour OR tumor)) AND (FDG OR Fludeoxyglucose OR fluorodeoxyglucose OR depreotide OR CGC OR PET OR SPECT OR NEOTECT OR NEOSPECT OR NEOTEC)
www.Clinicaltrials.gov
(FDG or PET or CT) | "Lung Neoplasms"
Theses OpenGrey
Lung cancer AND (FDG or PET or CT)

CONTRIBUTIONS OF AUTHORS

MSH conceived the idea for the review and wrote the protocol with input from DRB. EC designed the search strategy and performed the search. MSH and DRB screened the search results and selected the papers for inclusion. MSH, DRB, and MRF appraised the included papers and performed the data extractions. JZ and VA devised the analysis strategy, and JZ conducted the analyses. All the authors contributed to the writing of the review and approved the final manuscript.

DECLARATIONS OF INTEREST

Mia Schmidt-Hansen: nothing to declare.

David R Baldwin: nothing to declare.

Elise Hasler: nothing to declare.

Javier Zamora: nothing to declare.

Víctor Abraira: nothing to declare.

Marta Roqué i Figuls: nothing to declare.

This review presents an extension to a systematic review undertaken as part of the 2011 NICE clinical guideline on 'The Diagnosis and Treatment of Lung Cancer' (update) (NICE 2011).

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We have edited and updated the Background section to reflect the new structure of this section in Review Manager 5.2 (Review Manager 2012).

In discussion with the Cochrane Lung Cancer Group's Trials Search Co-ordinator, it was felt that a comprehensive search in MEDLINE and Embase supplemented with a careful review of reference lists of relevant papers combined with citation tracking of the included studies, rather than the long list of databases stated in the protocol, would be appropriate to identify relevant studies for this review. The search strategy detailed in the protocol was amended to become more refined and focused, and the 'staging' terms were removed as the inclusion of those terms weakened the proposed overall search strategy.

Regarding prespecified cut-off values for PET-CT positivity, we selected this item for preplanned sensitivity analyses to assess whether the results were sensitive to whether the cut-off values for test positivity were specified a priori or posthoc. However, on appraising the included studies, it became apparent that this item does not apply to at least half of the included studies, that is, the studies that do not use an explicitly quantitative test measure (i.e., SUV). Because when no quantitative criterion has been employed, the answer to this item is no without this in itself giving rise to a problem. We therefore decided to just incorporate this potential source of bias into

the 'Risk of bias' assessment for the index test and to limit the assessment of the influence of this item to the sensitivity analysis of the risk of bias for the index test.

We performed the heterogeneity analyses using the bivariate model instead of the HSROC model, because on reflection, we felt this approach made it easier to interpret the effect of a covariate on the accuracy indices and thereby, the results clearer.

INDEX TERMS

Medical Subject Headings (MeSH)

Carcinoma, Non-Small-Cell Lung [pathology; radiography; *radionuclide imaging]; Cross-Sectional Studies; Lung Neoplasms [pathology; radiography; *radionuclide imaging]; Lymph Nodes [pathology; *radionuclide imaging]; Lymphatic Metastasis; Mediastinum [radionuclide imaging]; Multimodal Imaging [*methods]; Positron-Emission Tomography [instrumentation; *methods]; Prospective Studies; Retrospective Studies; Tomography, X-Ray Computed [instrumentation; methods]

MeSH check words

Humans