HTA REPORT

FDG-PET/CT for cancer staging

APPENDICES

September 2012

Regione Umbria | Regione Puglia | Regione Siciliana | Regione Lazio | Provincia Autonoma Trento | Regione Emilia Romagna

Index

APPENDIX 1

FDG-PET/CT for staging of brain tumours

HTA report - KCE 2009

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. "For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients; • Absence of adequate reference standard; • Absence of patient-based analysis; • Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard). Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the

	performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	The 2009 KCE report (KCE 2009) conclusions for brain cancer staging is based on the AHRQ 2008 report. It identified two studies using PET to stage patients with suspected primary glioma and one study to stage patients with primary astrocytomas. All studies used histology/biopsy as reference standard. In the 2 studies sensitivity was only 63% and 75% and specificity was 100% and 0%. KCE report concluded that FDG-PET scanning is insufficiently accurate to be recommended for staging of brain cancer.
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

Characteristics of excluded studies

De Wever 2010

Reason for exclusion	No primary tumor

Dunet 2010

Reason for exclusion	No FDG

Giovacchini 2009

Reason for exclusion	No PET/CT

Jora 2011

Reason for exclusion	No staging

Li 2012

Reason for exclusion	No staging

1 CDSR, DARE, HTA database, CENTRAL search strategy

"Positron-Emission Tomography" [MeSH descriptor explode all trees]
OR
"Fluorodeoxyglucose F18" [MeSH descriptor explode all trees]
OR.
"positron emission tomography": ti,ab,kw
OR.
pet*: ti,ab,kw
OR.
pet scan*: ti,ab,kw
OR.
"Fluorodeoxyglucose F18": ti,ab,kw
OR.
fdg NEAR/2 18: ti,ab,kw
AND
Brain Neoplasms [Mesh explodes all trees]
OR
Intracranial Neoplasm*: ti,ab,kw
OR
Intracranial Cancer*: ti,ab,kw
OR
Brain Cancer*: ti,ab,kw
OR
Brain Neoplasm*:ti,ab,kw
OR
Brain Tumor*:ti,ab,kw
OR
Intracranial Tumor*:ti,ab,kw
OR
Brain NEAR/4 cancer*: ti,ab,kw
OR
Brain NEAR/4 neoplasm*: ti,ab,kw
OR
Intracranial NEAR/4 cancer*: ti,ab,kw
OR
Intracranial NEAR/4 neoplasm*: ti,ab,kw
OR
Intracranial NEAR/4 tumor*: ti,ab,kw
OR
Brain NEAR/4 tumor*: ti,ab,kw

2 MEDLINE search strategy

"Fluorodeoxyglucose F18"[Mesh] OR
"2-Fluoro-2-deoxyglucose" [All Fields] OR
"18F Fluorodeoxyglucose" [All Fields] OR
"F 18 Fluorodeoxyglucose" [All Fields] OR
Fludeoxyglucose* [All Fields] OR
"2 fluoro 2 deoxy d glucose"[All Fields] OR
18fluorodesoxyglucose*[All Fields] OR
fluorodeoxyglucose*[All Fields] OR
"fluorine 18 fluorodeoxyglucose" [All Fields] OR
18f dg*[All Fields]) OR
18fluorodeoxyglucose*[All Fields] OR
18fdg [All Fields] OR
18 fdg* [All Fields] OR
fdg 18* [All Fields] OR
fdg/* [All Fields] OR
"fdg pet"[All Fields] OR

“Positron-Emission Tomography”[Mesh] OR
“positron emission tomography” [title/abstract] OR
pet [title/abstract] OR
“pet scan” [All Fields] OR
“pet scans” [All Fields] OR
“pet scanner” [All Fields] OR
petscan [All Fields]

AND

Brain Neoplasms [Mesh explodes all trees] OR
“Brain Neoplasm” [Title/Abstract] OR
“Brain Neoplasms” [Title/Abstract] OR
“Intracranial Neoplasm” [Title/Abstract] OR
“Intracranial Neoplasms” [Title/Abstract] OR
“Brain Cancer” [Title/Abstract] OR
“Brain Cancers” [Title/Abstract] OR
“Intracranial Cancer” [Title/Abstract] OR
“Intracranial Cancers” [Title/Abstract] OR
“brain tumor” [Title/Abstract] OR
“intracranial tumor” [Title/Abstract]

Limits:from Genuary 2009; humans

3 EMBASE search strategy

“positron emission tomography”/syn OR
“fluorodeoxyglucose f 18”/exp OR “fluorodeoxyglucose f 18”/syn OR
“computer assisted emission tomography”/exp OR “computer assisted emission tomography” OR
pet OR
“pet scans” OR
“pet scanner” OR
“pet scan” OR
“pet/ct scan” OR
“pet/ct scans” OR
“pet/ct” OR
OR“positron emission tomography/computed tomography” OR
OR pet NEAR/4 scan*
OR pet NEAR/4 ct

AND

“Brain Neoplasms”/de, syn, Keyword OR
“Brain Neoplasms”/exp OR
“brain cancer”/de, syn, Keyword” OR
“brain cancers”/de, syn, Keyword” OR
“intracranial Neoplasms”/de, syn, Keyword OR
“Intracranial cancer”/de, syn, Keyword OR
“Intracranial cancers”/de,syn;keyword OR
“Intracranial neoplasm”/de,syn, keyword OR
“brain tumor“/de,syn, keyword OR
“brain tumors”/de,syn, keyword OR
“brain cancer”: ti, ab. OR
“brain neoplasm”: ab:ti OR
“brain neoplasms” : ab:ti OR
“brain cancers”: :ab:ti OR
“brain tumor”: :ab:ti OR

“brain tumors” :ab:ti OR
“intracranial cancers”: :ab:ti OR
“intracranial cancer”:ab:ti OR
“intracranial neoplasm”:ab:ti OR
“Intracranial neoplasms”:ab:ti OR
“intracranial tumor”: ab:ti OR
“intracranial tumors”:ab:ti OR
brain NEAR/4 neoplasm* OR
brain NEAR/4 cancer* OR
intracranial NEAR/4 neoplasm* OR
intracranial NEAR/4 cancer* OR
Brain NEAR/4 tumor* OR
Intracranial NEAR/4 tumor*
Limits:from Genuary 2009; humans; “article” OR “review”/it OR “short survey

APPENDIX 2

FDG-PET/CT for staging of head and neck cancer

Characteristics of included studies

HTA report - ASSR 2012 head and neck cancer

Document ID	ASSR 2012 head and neck cancer
Objectives	to define criteria for appropriate use of FDG-PET for patients with head and neck cancer
Methods	A panel of experts working in Health Trusts and Teaching Hospitals of Emilia-Romagna was convened to discuss and agree on the methodology for a research programme aimed at defining the criteria for appropriate use of PET in head and neck cancer. On the basis of the clinical pathway of patients with head and neck cancer the panel examined and assessed the role of FDG-PET for 9 clinical indications (diagnosis of head and neck cancer; detection of unknown primary head and neck cancer in patients with metastatic cervical lymph nodes; N staging of patients with head and neck cancer; M staging and detection of synchronous second primary tumor in patients with locally advanced head and neck cancer; target volume definition of curative radiation treatment; evaluation of early response to neo-adjuvant/induction therapy; evaluation of response to chemotherapy or radiotherapy at the end of treatment; follow up in patients with no suspicion of recurrence; diagnosis and staging of suspect distant recurrence). The following databases were searched for the period between January 2006 and March 2011: Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE - The Cochrane Library); Health Technology Assessment Database; Cochrane Central Register of Controlled Trials; National Library of Medicine's Medline database (PubMed); Elsevier's Embase. Language restrictions: English, Italian, French and Spanish. Selection criteria Type of studies: systematic reviews, RCTs, CCTs, cross-sectional diagnostic studies, prospective or retrospective cohort studies, case series of at least 10 patients Participants: patients with head and neck cancer Intervention: FDG-PET or CT/PET Reference standard: histology or clinical follow up Comparator: any other imaging technique Outcomes> sensitivity, specificity, LR, metabolic/tumor response, time to recurrence, local, local-regional and distant recurrence, disease free survival, disease survival, overall survival Assessment of methodological quality of studies The following criteria have been used for the quality assessment of different study designs.

	Systematic reviews: criteria drawn from the AMSTAR checklist Diagnostic cross sectional studies: criteria drawn from the QUADAS checklist Randomized controlled trials: criteria suggested by the Cochrane Handbook Case control studies and cohort studies: criteria drawn from the New Castle-Ottawa checklist Level of evidence for estimates of diagnostic accuracy were assigned according to GRADE categorization of the quality of evidence Each member of the panel voted the level of appropriateness for each clinical question. Two rounds of votes were requested for the judgment of appropriateness and results were analysed using the RAND/UCLA Appropriateness Method. The use of FDG-PET for a specific clinical indication was judged was judged as *appropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 7-9 score region as *inappropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 1-3 score region. Finally the use of FDG-PET was judged as *uncertain* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 4-6 score region or when no agreement was reached after the second round of voting. Clinical indications for which the panel does not reach an agreement on level of appropriateness after two rounds of voting also fall in the *uncertain* category.;
Conclusions	N STAGING OF PATIENTS WITH HEAD AND NECK CANCER - APPROPRIATE Use of FDG-PET for N staging of patients with primary head and neck cancer and with unclear results with conventional imaging (CT, MRI, ultrasound) has been judged appropriate by the panel during the first round of voting. Level of evidence for diagnostic accuracy of FDG-PET has been judged moderate, with estimates for sensitivity and specificity slightly higher than those of conventional imaging. Outcomes for patients correctly upstaged (true positives) have been voted “critical” (median score of 8, range 6-9), highlighting the importance attributed to the identification of node positive patients missed by conventional imaging. Consequences for patients testing negative (true and false negatives) and for false positives have also been judged critical, though with a lower median score and much wider range of votes. M STAGING OF PATIENTS AND DETECTION OF SYNCHRONOUS SECOND PRIMARY TUMOR IN PATIENTS WITH LOCALLY ADVANCED HEAD AND NECK CANCER - APPROPRIATE At the first voting round the panel agreed to judge appropriate the use of FDG-PET for M staging of advanced head and neck cancer in patients with negative or equivocal results from conventional imaging. Level of evidence for diagnostic accuracy of FDG-PET was judged moderate with estimates for sensitivity higher than conventional imaging. All clinical outcomes were considered “critical” (median score 8), with a closer range (between 7 and 8) for patients correctly upstaged, highlighting the added value of FDGPET in identifying patients with distant metastases or second primary tumors missed by conventional imaging.
Notes	Meta-analysis of diagnostic accuracy estimates was not performed

PS - N staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Liao 2011	FDG-PET/CT	473	oropharyngeal or hypopharyngeal squamous cell carcinoma 57% clinically negative neck patients	77.7	58.0

PS - Liao 2011

Clinical features and settings	oral cavity squamous cell carcinoma; Country: Taiwan
Participants	Patients with untreated oropharyngeal or hypopharyngeal squamous cell carcinoma 473 patients (445 males) PET-negative PET-positive (n = 199) (n=274) Male 185 (93.0) 260 (94.9) Female 14 (7.0) 14 (5.1) Age (y) <=40 38(19.1) 52(19.0) >40 161 (80.9) 222 (81.0) Clinical N status cN0 176 (88.4) 92 (33.6) cN+ 23(11.6) 182(66.4) Pathologic T status pT1–2 143(71.9) 132 (48.2) pT3–4 56(28.1) 142 (51.8) Pathologic N status pN0 152 (76.4) 110 (40.1) pN+ 47 (23.6) 164 (59.9) Level IV or V metastases

	No 197 (99.0) 261 (95.3) Yes 2(1.0) 13(4.7)	
Study design	diagnostic cross sectional study with prospective recruitment	
Target condition and reference standard(s)	neck lymph node metastases Reference standard: pathologic results at neck dissection	
Index and comparator tests	FDG-PET/CT; comparator: none	
Follow-up	not applicable	
Notes		

Assessment of methodological quality table

Item	**Authors' judgement**	**Support for judgement**
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Yes (cutoff of 2 for the 18F-FDG uptake score) Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

SR – M staging - Xu 2011a

Disease	head and neck cancer
Index test	FDG-PET/CT
Comparators	none
Reference standard	Histopathologic analysis or clinical and imaging follow-up for at least 6 months
Target	diagnostic accuracy for initial M staging (including second primary cancer)
Studies included	diagnostic accuracy primary studies with prospective or retrospective set-up of study
Years covered by the search	up to March 2011
Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE, EMBASE, EBM review databases)
Characteristics of included studies clearly reported in tables	Partially (not reported clinical features of included participants)
Methodological quality of primary studies assessed; criteria reported	Yes (study design and QUADAS tool)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	8 studies
Design of included studies	4 studies prospective design, 4 studies retrospective design
N. of included patients	824
Diagnostic accuracy results (with heterogeneity)	All patients FDG-PET/CT sensitivity 88.2% (95% CI 79.8-93.9%) specificity 95.1% (95% CI 93.2-96.5%)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? 4 studies out of 8 prospective design Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? All but 1 study with visual analysis of images Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	High risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? No (all studies) Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Unclear risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - M staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Chan 2011	FDG-PET/CT	103	patients with oropharyngeal or hypopharyngeal carcinoma, mainly locally advanced (about 75% of included patients)	83.3 (95% CI 58.6–96.4)	95.3 (95% CI 88.4–98.7)
	WB-MRI			66.7 (95% CI 41.0–86.7)	96.5 (95% CI 90.0–99.3)

PS - Chan 2011

Clinical features and settings	oropharyngeal and hypopharyngeal carcinoma; Country: Taiwan
Participants	Patients with untreated oropharyngeal or hypopharyngeal squamous cell carcinoma 103 patients (97 males); mean age (SD) 53.6 ± 9 Oropharynx 54 (52.4%); hypopharynx 49 (47.6%) T stage 1: 15 (14.6%) T stage 2: 24 (23.3%) T stage 3: 11 (10.7%) T stage 4: 53 (51.4%) N stage 0: 19 (18.4%) N stage 1: 5 (4.9%) N stage 2: 65 (63.1%) N stage 3: 14 (13.6%)
Study design	diagnostic cross sectional study with prospective recruitment

Target condition and reference standard(s)	distant metastases or second primary cancer Reference standard: pathological proof or evidence of progression at follow-up
Index and comparator tests	FDG-PET/CT; comparator: 3.0 Tesla Whole Body MRI (WB-MRI)
Follow-up	12 months
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Visual interpretation Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Low risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Yes Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? No: 13 patients excluded Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Any staging

Author, year	Number of participants	Population	Follow up	Outcome	FDG-PET/CT arm (% of participants)	no FDG-PET/CT arm (% of participants)	p*
Fried 2012	116 (58 FDG-PET/CT arm; 58 no FDG-PET/CT arm)	head and neck squamous cell carcinoma	24 months	Local control Regional control Distant metastases Cause-specific survival Overall survival	75.6 (95% CI 65–88) 81.0 (95% CI 72–93) 82.4 (95% CI 72–94) 70.8 (95% CI 59–85) 68.1 (95% CI 56–83)	70.1 (95% CI 58–84 0,66) 76.0 (95% CI 65–89) 84.6 (95% CI 75–96) 66.4 (95% CI 54–82) 63.5 (95% CI 51–79	0.66 0.26 0.65 0.55 0.57

*On multivariate analysis pretreatment PET imaging did not influence any endpoint

PS - Fried 2012

Clinical features and settings	head and neck squamous cell carcinoma; Country: USA
Participants	From a retrospective chart review, 249 patients that received definitive radiotherapy alone or chemoradiotherapy during the FDG-PET era from March 2002 to February 2010 were retrieved. 100 patients (40%) had a pretreatment FDG-PET for staging. Patients who had FDG-PET (PET cohort) were matched to those who did not (No PET cohort). From this matching process 116 patients were identified, 58 in each cohort. Patients were matched for T classification, N classification (according to CT), primary site (nasopharynx, oral cavity, oropharynx, larynx, or hypopharynx), and smoking status. PET cohort N / No PET cohort N Hypopharynx 5 (9%) / 5 (9%) Larynx 10 (17%) / 10 (17%) Nasopharynx 2 (3%) / 2 (3%) Oral cavity 3 (5%) / 3 (5%) Oropharynx 38 (66%) / 38 (66%)

	T1	6 (10%)	6 (10%)
	T2	14 (24%)	14 (24%)
	T3	18 (31%)	18 (31%)
	T4	20 (34%)	20 (34%)
	N0	5 (9%)	5 (9%)
	N1	5 (9%)	5 (9%)
	N2	35 (60%)	1 (60%)
	N3	13 (22%)	13 (22%)
	Median age	57 years	55 years
	Radiotherapy		
	Conventional	17 (29%)	24 (41%)
	IMRT	41 (71%)	34 (59%)
	Male	45 (76%)	49 (85%)
	Female	13 (24%)	9 (15%)
	Chemotherapy		
	Yes	55 (95%)	57 (98%)
	No	3 (5%)	1 (2%)
Study design	retrospective matched cohort study		
Target condition and reference standard(s)	local control, regional control, freedom from distant metastasis, cause-specific survival, overall survival Reference standard: not applicable		
Index and comparator tests	FDG-PET/CT or FDG-PET for staging; comparator: staging without FDG-PET		
Follow-up	24 months		
Notes			

Assessment of methodological quality table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	retrospective matched cohort study
Allocation concealment (selection bias)	High risk	
Blinding of participants and personnel (performance bias)	High risk	open study
Blinding of outcome assessment (detection bias)	Unclear risk	
Incomplete outcome data (attrition bias)	Low risk	
Selective reporting (reporting bias)	Low risk	
Other bias	Low risk	

Characteristics of excluded studies

Boktor 2012

Reason for exclusion	retrospective study

Davis 2011

Reason for exclusion	abstract at congress

de Casso 2012

Reason for exclusion	retrospective study

Ghanooni 2011

Reason for exclusion	study on treatment response evaluation

Guden 2010

Reason for exclusion	retrospective study

Gupta 2011

Reason for exclusion	abstract at congress

Haerle 2011

Reason for exclusion	retrospective study

Haerle 2011a

Reason for exclusion	per-lesion analysis

Haerle 2011b

Reason for exclusion	retrospective study

Huang 2011

Reason for exclusion	per-lesion analysis

Kim 2011

Reason for exclusion	abstract at congress

Kim 2011a

Reason for exclusion	retrospective study

Kondo 2011

Reason for exclusion	retrospective study

Krabbe 2011

Reason for exclusion	retrospective study

Kurien 2011

Reason for exclusion	not staging

Lakshmipathy 2011

Reason for exclusion	abstract at congress

Lamarre 2011

Reason for exclusion	retrospective study

Law 2011

Reason for exclusion	prognostic study

Nakaminato 2012

Reason for exclusion	study on first diagnosis

Ng 2011

Reason for exclusion	study on recurrrence

Nguyen 2011

Reason for exclusion	retrospective study

Oh 2011

Reason for exclusion	abstract at congress

Pietka 2011

Reason for exclusion	full-text not found

Radhakrishnan 2012

Reason for exclusion	not head and neck cancer

Wu 2011

Reason for exclusion	not N or M staging

Xie 2011

Reason for exclusion	systematic review on prognosis

Xu 2011b

Reason for exclusion	old version of included systematic review

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”: ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw or
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. “Head and Neck Neoplasms” [MeSH descriptor explode all trees]
10. 8 AND 9

Publication date: January 2011 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18” [Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg-pet”[All Fields]
17. “Positron-Emission Tomography” [Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. “head and neck cancer”[Title/Abstract]
26. head and neck cancers”[Title/Abstract]
27. “head and neck neoplasms”[Title/Abstract])
28. “head and neck neoplasm”[Title/Abstract])
29. “oral neoplasm”[Title/Abstract]
30. “oral neoplasms”[Title/Abstract]
31. “cancer of mouth”[Title/Abstract]
32. “oral cancer”[Title/Abstract]
33. “oral cancers”[Title/Abstract]
34. “gingival neoplasms”[Title/Abstract]
35. “gingival neoplasm”[Title/Abstract]
36. “congenital epulides”[Title/Abstract]
37. “congenital epulis”[Title/Abstract]
38. leukokeratoses[Title/Abstract]
39. leukokeratosis[Title/Abstract]
40. leukoplakia[Title/Abstract]

41. leukoplakias[Title/Abstract
42. “lip cancer”[Title/Abstract]
43. “lip cancers”[Title/Abstract]
44. “lip neoplasms”[Title/Abstract]
45. “lip neoplasm”[Title/Abstract]
46. “palatal neoplasm” [Title/Abstract]
47. “palatal neoplasms”[Title/Abstract]
48. “salivary gland neoplasm”[Title/Abstract]
49. “salivary gland neoplasms”[Title/Abstract]
50. “salivary gland cancer”[Title/Abstract]
51. “salivary gland cancers”[Title/Abstract]
52. “parotid neoplasm”[Title/Abstract]
53. “parotid neoplasms”[Title/Abstract]
54. “parotid cancer”[Title/Abstract]
55. “parotid cancers”[Title/Abstract]
56. “parotid carcinomas”[Title/Abstract]
57. “gland neoplasm”[Title/Abstract]
58. “gland neoplasms”[Title/Abstract]
59. “tongue neoplasm”[Title/Abstract]
60. “tongue neoplasms”[Title/Abstract]
61. “tongue cancer”[Title/Abstract]
62. “tongue cancers”[Title/Abstract]
63. “otorhinolaryngological neoplasms”[Title/Abstract]
64. “otorhinolaryngological neoplasm”[Title/Abstract]
65. “otorhinolaryngological cancer”[Title/Abstract]
66. “otorhinolaryngological cancers”[Title/Abstract]
67. “auricular cancer”[Title/Abstract]
68. “auricular cancers”[Title/Abstract]
69. “auricular carcinoma”[Title/Abstract]
70. “ear neoplasm”[Title/Abstract]
71. “ear neoplasms”[Title/Abstract]
72. “ear cancer”[Title/Abstract]
73. “ear cancers”[Title/Abstract]
74. “laryngeal neoplasm”[Title/Abstract]
75. “laryngeal neoplasms”[Title/Abstract]
76. “laryngeal cancer”[Title/Abstract]
77. “laryngeal cancers”[Title/Abstract]
78. “larynx neoplasm”[Title/Abstract]
79. “larynx neoplasms”[Title/Abstract]
80. “larynx cancer”[Title/Abstract]
81. “larynx cancers”[Title/Abstract]
82. “nose neoplasms”[Title/Abstract]
83. “nose neoplasm”[Title/Abstract]
84. “nose cancer”[Title/Abstract]
85. “nose cancers”[Title/Abstract]
86. “sinus neoplasm”[Title/Abstract]
87. “sinus neoplasms”[Title/Abstract]
88. “paranasal sinus cancer”[Title/Abstract]
89. “paranasal sinus cancers”[Title/Abstract]
90. “sinus cancer”[Title/Abstract]
91. “sinus cancers”[Title/Abstract]
92. “pharyngeal neoplasm”[Title/Abstract]
93. “pharyngeal neoplasms”[Title/Abstract]
94. “pharyngeal cancer”[Title/Abstract]
95. “pharyngeal cancers”[Title/Abstract]
96. “pharynx cancer”[Title/Abstract]
97. “pharynx cancers”[Title/Abstract]
98. “hypopharyngeal cancer”[Title/Abstract]
99. “hypopharyngeal cancers”[Title/Abstract]
100. “nasopharynx cancer”[Title/Abstract]
101. “nasopharynx cancers”[Title/Abstract]
102. “oropharyngeal neoplasm”[Title/Abstract]

103. oropharyngeal neoplasms”[Title/Abstract]
104. “oropharyngeal cancer”[Title/Abstract]
105. “oropharyngeal cancers”[Title/Abstract]
106. “oropharynx cancer”[Title/Abstract]
107. “oropharynx cancers”[Title/Abstract]
108. “tonsil cancer”[Title/Abstract]
109. “tonsil cancers”[Title/Abstract]
110. “tonsillar neoplasm”[Title/Abstract
111. “tonsillar neoplasms”[Title/Abstract
112. “tonsillar cancer”[Title/Abstract
113. “tonsillar cancers”[Title/Abstract
114. “Mouth Neoplasms”[Mesh]
115. “Head and Neck Neoplasms”[Mesh:noexp]
116. “Otorhinolaryngologic Neoplasms”[Mesh]
117. 25/116 OR
118. 24 AND 117
Limit: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2011 - March 2012

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “positron emission tomography”/exp
3. “fluorodeoxyglucose f 18”/exp
4. “fluorodeoxyglucose f 18”/syn
5. “computer assisted emission tomography”/exp
6. “computer assisted emission tomography”/tw
7. pet/tw
8. “pet scans”/tw
9. “pet scanner”/tw
10. “pet scan”/tw
11. “pet/ct scan”/tw
12. “pet/ct scans”/tw
13. “pet/ct”/tw
14. “positron emission tomography/computed tomography”/tw
15. pet NEAR/4 scan*
16. pet NEAR/4 ct
17. 1/15 OR
18. “head and neck cancer”/exp
19. “head and neck cancer”/syn
20. “head and neck cancer”/tw
21. “head cancer”/de
22. “head cancer”/tw
23. “nose cancer”/exp
24. “nose cancer”/tw
25. “lip cancer”/exp
26. “lip cancer”/tw
27. “mouth cancer”/exp
28. “mouth cancer”/syn
29. mouth cancer”/tw
30. “neck cancer”/exp
31. “neck cancer”/tw
32. “paranasal sinus cancer”/tw
33. “paranasal sinus cancer”/exp
34. “pharynx cancer”/exp
35. “pharynx cancer”/tw
36. “salivary gland cancer”/exp
37. “salivary gland cancer”/tw
38. “tongue cancer”/exp
39. “tongue cancer”/tw
40. “tonsil cancer”/exp

41. “tonsil cancer”/tw
42. 18/42 OR
43. 17 AND 42
Limit: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2011 - March 2012

APPENDIX 3

FDG-PET/CT for staging of non-small cell lung cancer

Characteristics of included studies

HTA report - ASSR 2012 lung cancer

Document ID	ASSR-RER 2012 - Lung cancer
Objectives	to define criteria for appropriate use of FDG-PET for patients with non-small cell lung cancer
Methods	A panel of experts working in Health Trusts and Teaching Hospitals of Emilia-Romagna was convened to discuss and agree on the methodology for a research programme aimed at defining the criteria for appropriate use of PET in non-small cell lung cancer. On the basis of the clinical pathway of patients with non-small cell lung cancer the panel examined and assessed the role of FDG-PET for 7 clinical indications (characterization of solitary pulmonary nodules ≥1 cm; staging of patients with non-small cell lung cancer (NSCLC); target volume definition of radiation treatment with curative intent in patients treated for lung cancer; during-treatment evaluation of response to neo-adjuvant therapy in patients treated for lung cancer - NSCLC; end of treatment evaluation of response to neo-adjuvant therapy in patients treated for lung cancer - NSCLC; follow up of patients treated for lung cancer with no suspicion of recurrence - NSCLC; diagnosis and staging of suspected loco-regional recurrence in patients treated for lung cancer - NSCLC). The following databases were searched for the period between January 2006 and September 2010: Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE - The Cochrane Library); Health Technology Assessment Database; Cochrane Central Register of Controlled Trials; National Library of Medicine's Medline database (PubMed); Elsevier's Embase. Language restrictions: English, Italian, French and Spanish. Selection criteria Type of studies: systematic reviews, RCTs, CCTs, cross-sectional diagnostic studies, prospective or retrospective cohort studies, case series of at least 10 patients Participants: patients with non-small cell lung cancer Intervention: FDG-PET or CT/PET Reference standard: histology or clinical follow up Comparator: any other imaging technique Outcomes> sensitivity, specificity, LR, metabolic/tumor response, time to recurrence, local, local-regional and distant recurrence, disease free survival, disease survival, overall survival Assessment of methodological quality of studies The following criteria have been used for the quality assessment of different study designs.

	Systematic reviews: criteria drawn from the AMSTAR checklist Diagnostic cross sectional studies: criteria drawn from the QUADAS checklist Randomized controlled trials: criteria suggested by the Cochrane Handbook Case control studies and cohort studies: criteria drawn from the New Castle-Ottawa checklist Level of evidence for estimates of diagnostic accuracy were assigned according to GRADE categorization of the quality of evidence Each member of the panel voted the level of appropriateness for each clinical question. Two rounds of votes were requested for the judgment of appropriateness and results were analysed using the RAND/UCLA Appropriateness Method. The use of FDG-PET for a specific clinical indication was judged was judged as *appropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 7-9 score region as *inappropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 1-3 score region. Finally the use of FDG-PET was judged as *uncertain* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 4-6 score region or when no agreement was reached after the second round of voting. Clinical indications for which the panel does not reach an agreement on level of appropriateness after two rounds of voting also fall in the *uncertain* category.;
Conclusions	STAGING OF PATIENTS WITH NON-SMALL CELL LUNG CANCER - NSCLC - APPROPRIATE The panel agreed at the first voting round that the use of FDG-PET as an add on test in NSCLC staging is appropriate. The level of evidence supporting this indication is moderate with FDG-PET performing well in identifying mediastinal involvement or distant metastases missed by CT. While avoiding unnecessary surgery (consequences for true positives) has been considered important, undergoing futile surgery (consequences for false negatives) or not undergoing a potentially curative radical surgery (consequences for false positives) have been considered critical outcomes with median scores of 8 and 7 respectively, confirming the need for thorough and accurate pre-treatment staging.
Notes	Meta-analysis of diagnostic accuracy estimates was not performed

SR - Lv 2011 - N staging

Disease	non-small cell lung cancer
Index test	FDG-PET/CT (7 studies had SUV cut-off of 2.5, 1 of 3.0, 1 of 5.2, 5 performed a visual qualitative analysis)
Comparators	none
Reference standard	histological examination of lymph nodes by surgery or biopsy
Target	diagnostic accuracy for mediastinal lymph node staging
Studies included	diagnostic accuracy primary studies with prospective or retrospective design
Years covered by the search	up to December 2010
Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE, EMBASE, SpringerLink)
Characteristics of included studies clearly reported in tables	Partially (not reported the stage of patients at inclusion)
Methodological quality of primary studies assessed; criteria reported	Yes (QUADAS tool)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	14 primary studies (11 studies using patient as unit of analysis; 9 studies using nodes as unit of analysis)
Design of included studies	Cross sectional diagnostic accuracy studies, with prospective (n. 7) or retrospective recruitment (n. 7)
N. of included patients	2550 (range 46-674); 2191 (range 51-674) in studies using patient as unit of analysis
Diagnostic accuracy results (with heterogeneity)	Per-patient based analysis pooled weighted sensitivity 76% (95% CI: 65–84%) pooled weighted specificity 88% (95% CI: 82–92%) Authors found significant heterogeneity in the pooled analysis. This result was reported to be not unexpected because the studies adopted different SUV cutoffs as diagnostic criteria

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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 (50% of studies with retrospective design) Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Yes Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear in 40% of studies Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Unclear Were all patients included in the analysis? Unclear

		Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

SR - Zhao 2011 - N staging

Disease	non-small cell lung cancer
Index test	FDG-PET/CT
Comparators	none
Reference standard	histological examination of lymph nodes by surgery or biopsy
Target	diagnostic accuracy for mediastinal lymph node staging
Studies included	diagnostic accuracy primary studies with prospective or retrospective design
Years covered by the search	up to July 2011
Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE, EMBASE, EBM review databases)
Characteristics of included studies clearly reported in tables	Partially (not reported the stage of patients at inclusion)
Methodological quality of primary studies assessed; criteria reported	Yes (QUADAS tool, however data reported only as final sum score without any detail)
Meta-analysis performed with appropriate statistic	Yes

methods (including heterogeneity evaluation)	
N. of included studies	20 primary studies (14 studies using patient as unit of analysis; 14 studies using nodes as unit of analysis)
Design of included studies	Cross sectional diagnostic accuracy studies, with prospective (n. 11) or retrospective recruitment (n. 9)
N. of included patients	3028; 2087 in studies using patient as unit of analysis
Diagnostic accuracy results (with heterogeneity)	Per-patient based analysis pooled sensitivity 72% (95% CI: 68–75%) pooled specificity 90% (95% CI: 88–91%)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Unclear Did patients receive the same reference standard? Unclear Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - mediastinal N staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Fischer 2009	conventional staging# and FDG-PET/CT	98	suspected potentially resectable NSCLC	75.0 (95% CI 59.0-86.0)	100 (95% CI 94.0-100)
	conventional staging# without FDG-PET/CT	91		59.0 (95% CI 41.0-74.0)	98.0 (95% CI 91.0-100)
Gunluoglu 2011	FDG-PET/CT	168	NSCLC patients suitable for thoracotomy	71.0 (95% CI 57—83)	75.0 (95% CI 66—82)
	mediastinoscopy			84.0 (95% CI 70—92)	100 (95% CI 96—100)
	ESTS' guidelines 2007*			84.0 (95% CI 70—92)	100 (95% CI 96—100)
Ohnishi 2011	FDG-PET/CT	120§	suspected potentially resectable NSCLC	47.4 (95% CI 32.5–62.7)	87.5 (95% CI 77.9–93.3)
	EBUS–TBNA			64.1 (95% CI 48.4–77.3)	100 (95% CI 94.9–100)
	EUS–FNA			48.7 (95% CI 33.9–63.8)	100 (95% CI 94.9–100)
	EBUS–TBNA or EUS–FNA			71.8 (95% CI 56.2–83.5)	100 (95% CI 94.9–100)
Ohno 2011	FDG-PET/CT (visual)	250	T1 or T2 NSCLC	69.9	91.7
	FDG-PET/CT (SUVmax)			74.2	92.4
	MRI STIR turbo SE imaging (visual)			77.4	88.5
	MRI STIR turbo SE imaging (lymph node–to-saline ratio)			82.8	89.2
	MRI STIR turbo SE imaging (lymph node–to-			82.8	89.2

	muscle ratio)				
	DW MR imaging (visual)			71.0	89.8
	DW MR imaging (apparent diffusion coefficient)			74.2	90.4
Sivrikoz 2012	FDG-PET/CT	68	suspected potentially resectable NSCLC	61.0 (95% CI 46–65)	98.0 (95% CI 92–99)
	mediastinoscopy			81.8 (95% CI 63–82)	100 (95% CI 96–100)

Footnotes

* direct thoracotomy without invasive staging in patients who are evaluated to have peripheral, clinical stage I (T1—2 and N0) tumours on CT and PET; mediastinoscopy in patients who have central tumours, or are clinical T3—4 stage NSCLC cases, or have an mediastinal lymph node larger than 1.5 cm on thorax CT images; mediastinoscopy where PET reveals a tumour with low FDG uptake, or mediastinal or hilar lymph nodes showing FDG uptake.

§ 110 participants included in the analysis

conventional staging: clinical data, initial CT scanning, bronchoscopy

PS - Fischer 2009

Clinical features and settings	non-small cell lung cancer; Country: Denmark
Participants	189 patients newly diagnosed or highly suspected for non-small cell lung-cancer, considered to have operable disease after conventional-staging procedures (i.e., medical history, physical examination, blood test, contrast-enhanced CT scan of the chest and upper abdomen, and bronchoscopy). After conventional staging, eligible patients were randomly assigned in a 1:1 ratio to PET–CT and conventional staging, followed by further invasive diagnostic procedures such as mediastinoscopy and endoscopic or endobronchial ultrasonography (the PET–CT group = 98 participants), or to conventional staging and invasive diagnostic procedures alone (the conventional-staging group = 91 participants). Eleven patients in the PET–CT group did not undergo PET–CT because of an unacceptably long waiting time for a scan or technical problems with the PET–CT equipment. One patient underwent PET–CT but declined all further staging procedures and surgery. Mediastinoscopy was performed in 89 patients in the PET–CT group (91%) and 88 in the conventional-staging group (97%) (P = 0.33). PET-CT group Conventional-staging group N N

	Age (yr) mean 63 64 range 42–80 38–80 Male sex (no.) 53 49 Female sex (no.) 45 42 ECOG performance status (no.) 0–1 93 86 2 1 1 Not available 4 4 Tumor Size on CT (mm) Mean 46.5 43.6 Range 10.0–110.0 15.0–130.0 TNM stage based on CT of thorax and abdomen (no.) IA 13 9 IB 17 13 IIA 0 0 IIB 5 7 IIIA 26 28 IIIB 32 32 IV 5 2 Mediastinoscopy (no.) Total 89 88 Stage N2 to N3 disease 9 12 Histologic features at operation (no) Squamous-cell carcinoma 22 22 Adenocarcinoma 30 29 Large-cell carcinoma 4 12 Bronchoalveolar carcinoma 0 1 NSCLC with no further specification 5 4 Other 2 2 Benign lung lesion 0 3
Study design	Open randomised controlled trial. Randomization was performed centrally with the use of a permuted-block design, stratified according to sex and recruiting center. Before a decision to operate was made, a consensus on the TNM stage was reached by a pulmonologist and a thoracic surgeon on the basis of all available information (clinical data, initial CT scanning, PET–CT imaging, bronchoscopy, mediastinoscopy, and if available, endoscopic ultrasonography with fine-needle aspiration or endobronchial ultrasonography). Mediastinoscopy and endoscopic or endobronchial ultrasonography served as the standard for preoperative assessment of mediastinal lymph nodes. All patients with stage I to stage IIB NSCLC were offered surgery. Patients with involvement of mediastinal lymph nodes or distant metastases (stage IIIA [N2] to stage IV) were considered to have inoperable disease and were offered chemotherapy with or without radiotherapy. Positive findings on PET–CT were further evaluated by biopsy or other imaging techniques (ultrasonography, radiography, or magnetic resonance

	imaging) at the discretion of the referring clinician. The study was closed after the inclusion of only 189 patients (expected 215 assigned to each group) because of slow accrual.
Target condition and reference standard(s)	N staging Clinical outcomes: the primary end point of the study was the frequency of futile thoracotomies (a benign lung lesion, pathologically proven mediastinal lymph-node involvement (stage IIIA [N2]), stage IIIB or IV disease, inoperable T3 or T4 disease, or recurrent disease or death from any cause within 1 year after randomization); median survival; death Reference standard: pathologist after thoracotomy served as the reference (N final). For patients in whom thoracotomy was not performed, N stage assigned by mediastinoscopy, EUS-FNA or EBUS-TBNA
Index and comparator tests	FDG-PET/CT (and conventional diagnostic tools for staging: clinical data, initial CT scanning, bronchoscopy); comparator: staging without FDG-PET/CT (clinical data, initial CT scanning, bronchoscopy)
Follow-up	12 months
Notes	No financial support was received from companies that make PET–CT scanners

Assessment of methodological quality table (diagnostic accuracy)

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Yes (standardized uptake value above 2.5 judged to be positive) Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE	Unclear risk	Describe the reference standard and how it was conducted and

STANDARD (risk of bias)		interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? No Did patients receive the same reference standard? Unclear Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Gunluoglu 2011

Clinical features and settings	non-small-cell lung cancer; Country: Turkey
Participants	168 patients (149 males); mean age 60 years (range 30-84). Seventy-eight (46.4%) patients had squamous cell carcinoma, 57 (34%) adenocarcinoma, 9 (5.4%) adenosquamous cell carcinoma, 7 (4.2%) pleomorphic carcinoma, 1 (0.5%) large cell carcinoma, 16 (9.5%) had NSCLC without further specification. Patients suitable for thoracotomy (T stages not specified).
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	Mediastinal lymph node metastasis Reference standard: histological examination of lymph nodes by surgery or biopsy (during mediastinoscopy)
Index and comparator tests	FDG-PET/CT; comparator: mediastinoscopy; preoperative lymph-node staging recommendations from guidelines of the European Society of Thoracic Surgeons (stepwise add-on process using CT, FDG-PET/CT, mediastinoscopy).
Follow-up	not applicable
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Visual interpretation Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Unclear Were all patients included in the analysis? No: 17 patients excluded Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Ohnishi 2011

Clinical features and settings	non-small-cell lung cancer; Country: Japan
Participants	120 patients (79 males; median age 69 years, range 40–85) with suspected potentially resectable non-small cell lung cancer. Tumor staging, n T1 60 T2 43 T3 1 Nodal staging, n N0 70 N1 12 N2 23 N3 15 PET–CT was performed in all patients. Among the patients, distant metastases or pleural dissemination was revealed in five patients. These five patients were judged to be unresectable, and endobronchial ultrasound-guided transbronchial needle aspiration (EBUS–TBNA) and transesophageal endoscopic ultrasound-guided fine-needle aspiration (EUS–FNA) for final N staging was avoided. Therefore, EBUS and EUS were performed in 115 patients. Among them, 16 patients had N2 involvement, 4 of whom received induction chemoradiotherapy followed by surgical resection. The other 12 patients with N2 disease and 10 patients with N3disease were treated by chemotherapy or chemoradiotherapy. The remaining 89 patients were diagnosed with N0/N1disease by EUS–FNA and EBUS–TBNA. However, as five of them refused surgery, a final diagnosis of lymph node involvement was not available, and they were excluded from the assessment of N staging. After surgery in a total of 84 patients,79 were diagnosed with non-small-cell lung cancer, with a histological diagnosis of adenocarcinoma (n 47), squamous cell carcinoma (n 19), adenosquamous cell carcinoma (n 2), bronchioalveolar carcinoma (n 5), large-cell carcinoma (n 2), mucoepidermoid carcinoma (n 1), and large cell neuroendocrine carcinoma (n 3). Five patients were diagnosed as having benign disease: atypical adenomatous hyperplasia (n 2), sarcoidosis (n 1), intrapulmonary lymph node (n=1), and pulmonary tuberculosis (n=1).
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	Mediastinal lymph node metastasis Reference standard: histological examination of lymph nodes by surgery or

	biopsy (with EUS–FNA or EBUS–TBNA)
Index and comparator tests	FDG-PET/CT; comparator: EUS–FNA or EBUS–TBNA.
Follow-up	not applicable
Notes	The 5 patients with a final diagnosis of benign diseases were also included in the assessment of N staging

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Yes (standardized uptake value above 3.0 judged to be positive) Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	High risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

		Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? No Did patients receive the same reference standard? No Were all patients included in the analysis? No: 10 patients excluded Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Ohno 2011

Clinical features and settings	non-small-cell lung cancer; Country: Japan
Participants	250 patients (136 males; mean age, 73 years; age range, 61–83 years) with T1 or T2 non-small cell lung cancer as evaluated on chest radiographs or CT images. The final diagnosis of lung cancer and N stage disease was based on pathologic findings in resected specimens, which showed that 218 patients had adenocarcinoma, 23 had squamous cell carcinoma, 6 had large cell carcinoma, and three had adenosquamous cell carcinoma. Stage N0 disease was detected in 157 patients, N1 disease was detected in 72, N2 disease was detected in 16, and N3 disease was detected in five
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and	Mediastinal lymph node metastasis

reference standard(s)	Reference standard: histological examination of lymph nodes by surgery or biopsy (with mediastinoscopy)
Index and comparator tests	FDG-PET/CT; comparator: MRI (short inversion time inversion recovery STIR turbo spin-echo SE; diffusion-weighted DW)
Follow-up	not applicable
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Yes (both for quantitative analysis and qualitative analysis) Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

		Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Sivrikoz 2011

Clinical features and settings	non-small-cell lung cancer; Country: Turkey
Participants	68 patients (60 males; mean age 60.36 +/- 1.01 years, range: 43–78 years) with suspected or pathologically proven, localized, clinically resectable non-small cell lung cancer. squamous carcinoma in 44 (64.7%) patients, adenocarcinoma in 20 (29%) patients and other types in 4 (6.3%) patients (adenosquamous carcinoma in one patient, large cell carcinoma in one, adenocystic carcinoma in one and mucoepidermoid carcinoma in one patient) Seven (10.2%) patients who underwent invasive mediastinal staging were excluded fromsubsequent surgery due to N2 disease (n 5) or N3 disease (n 2). Sixty-one (89.3%) patients underwent invasive mediastinal staging followed by a thoracotomy during the same surgical session (systematic sampling of lymph nodes by thoracotomy and pulmonary resection if possible) due to nonmetastatic mediastinal lymph node N0 (n 54) or N1 (n 7) disease.

Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	Mediastinal lymph node metastasis Reference standard: histological examination of lymph nodes by surgery or biopsy (with mediastinoscopy)
Index and comparator tests	FDG-PET/CT; comparator: mediastinoscopy
Follow-up	not applicable
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Yes (both for quantitative analysis and qualitative analysis) Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer

(risk of bias)		to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

SR - Chang 2012 - M staging

Disease	lung cancer (majority on non-small cell lung cancer)
Index test	FDG-PET or FDG-PET/CT
Comparators	bone scintigraphy
Reference standard	histopathologic analysis and/or close clinical and imaging follow-up and/or radiographic confirmation by multiple imaging modalities
Target	diagnostic accuracy for bone metastasis
Studies included	diagnostic accuracy primary studies with prospective or retrospective design (not clear if studies on staging or recurrence or both)
Years covered by the search	up to August 2010
Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE, EMBASE, Cochrane Library)
Characteristics of included studies clearly reported in tables	Partially (not reported neither the clinical phase - staging / recurrence - nor the histological type of cancer)
Methodological quality of primary studies assessed; criteria reported	Yes (criteria recommended by Cochrane Methods Working Group on Systematic Reviews of Screening and Diagnostic Tests)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	7 primary studies (6 studies using patient as unit of analysis; 3 studies using lesion as unit of analysis)
Design of included studies	Cross sectional diagnostic accuracy studies, with prospective (n. 1) or retrospective recruitment (n. 6) and with direct comparison of FDG-PET or FDG-PET/CT with bone scintigraphy
N. of included patients	1794 (range 48-1000); 1746 (range 82-1000) in studies using patient as unit of analysis
Diagnostic accuracy results (with heterogeneity)	Per-patient based analysis FDG-PET or FDG-PET/CT pooled sensitivity 93% (95% CI 88–96%) (heterogeneity test P = 0.932) pooled specificity 95% (95% CI: 91–98%) (heterogeneity test P < 0.001)

	bone scintigraphy pooled sensitivity 87% (95% CI 79–93%) (heterogeneity test P = 0.06) pooled specificity 82% (95% CI: 62–92%) (heterogeneity test P < 0.001)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	High risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? 3 studies out of 7 with consecutive enrollment Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear (6 studies out of 7 with retrospective design) Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? 4 studies out of 7 If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	High risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? No in all studies Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s)

		and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Unclear Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	High risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? Unclear the clinical phase of patients (staging or suspected recurrence or both?); unclear the histological type of cancer; all studies consider bone metastasis only CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

SR - Liu 2011 - M staging

Disease	non-small cell lung cancer
Index test	FDG-PET or FDG-PET/CT
Comparators	bone scintigraphy, MRI
Reference standard	histopathological analysis and/or close clinical and imaging follow-up and/or radiographic confirmation by multiple imaging modalities
Target	diagnostic accuracy for bone metastasis
Studies included	diagnostic accuracy primary studies with prospective or retrospective design (not clear if studies on staging or recurrence or both)
Years covered by the search	up to January 2010

Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE, EMBASE)
Characteristics of included studies clearly reported in tables	Partially (not reported neither the clinical phase - staging / recurrence- nor the histological type of cancer)
Methodological quality of primary studies assessed; criteria reported	Yes (QUADAS tool)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	34 primary studies assessing FDG-PET or FDG-PET/CT (22 studies using patient as unit of analysis; 12 studies using lesion as unit of analysis); 5 primary studies assessing FDG-PET/CT (patient as unit of analysis); 11 primary studies assessing bone scintigraphy (patient as unit of analysis); 3 primary studies assessing MRI (patient as unit of analysis);
Design of included studies	Cross sectional diagnostic accuracy studies (not reported how many with prospective or retrospective design)
N. of included patients	5676; among studies using patient as unit of analysis, studies assessing FDG-PET or FDG-PET/CT included a total of 2446 patients, studies assessing bone scintigraphy included a total of 1537 patients, studies assessing MRI included a total of 258 patients
Diagnostic accuracy results (with heterogeneity)	Per-patient based analysis FDG-PET/CT (5 studies) pooled sensitivity 94.6% (95% CI 91.1–97.0%) pooled specificity 97.5% (95% CI 96.6–98.3%) bone scintigraphy (11 studies) pooled sensitivity 91.8% (95% CI 89.1–94.1%) pooled specificity 68.8% (95% CI 65.8–71.6%) MRI (3 studies) pooled sensitivity 80.0% (95% CI 67.0–89.6%) pooled specificity 90.6% (95% CI 85.8–94.3%)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	High risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? No in all studies Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Unclear Were all patients included in the analysis? Unclear

		Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	High risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? Unclear the clinical phase of patients (staging or suspected recurrence or both?); unclear the histological type of cancer; all studies consider bone metastasis only CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

SR - Qu 2011 - M staging

Disease	non-small cell lung cancer and other type of lung cancer
Index test	FDG-PET or FDG-PET/CT
Comparators	bone scintigraphy, MRI
Reference standard	histopathology alone, histopathology plus clinical follow-up or clinical follow-up alone
Target	diagnostic accuracy for bone metastasis
Studies included	diagnostic accuracy primary studies with prospective or retrospective design (not clear if studies on staging or recurrence or both)
Years covered by the search	up to August 2010
Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE, EMBASE, Cochrane Library)
Characteristics of included studies clearly reported in tables	Partially (not reported neither the clinical phase - staging / recurrence- nor the histological type of cancer)
Methodological quality of primary studies assessed; criteria reported	Yes (QUADAS tool)

Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	17 primary studies (11 including only NSCLC patients; 5 including both NSCLC and SCLC patients; 1 including only SCLC patients) 7 studies of FDG-PET/CT using patient as unit of analysis; 2 studies of FDG-PET/CT using lesion as unit of analysis; 5 studies of FDG-PET using patient as unit of analysis; 5 studies of FDG-PET/CT using lesion as unit of analysis; 3 studies of MRI using patient as unit of analysis; 3 studies of MRI using lesion as unit of analysis; 12 studies of bone scintigraphy using patient as unit of analysis; 4 studies of bone scintigraphy using lesion as unit of analysis;
Design of included studies	Cross sectional diagnostic accuracy studies (5 studies out of 17 with prospective design; the other studies with retrospective or unclear design)
N. of included patients	2940; Considering studies using patient as unit of analysis: studies assessing FDG-PET included a total of 840 patients, studies assessing FDG-PET/CT 1855 patients (only non-small cell lung cancer:1644), studies assessing bone scintigraphy 2285 patients (only non-small cell lung cancer:1640), studies assessing MRI 252 patients
Diagnostic accuracy results (with heterogeneity)	Per-patient based analysis FDG-PET/CT (non-small cell lung cancer) pooled sensitivity 92.0% (95% CI 88.0–95.0%) pooled specificity 98.0% (95% CI 97.0–99.0%) bone scintigraphy (non-small cell lung cancer) pooled sensitivity 85.0% (95% CI 80.0–89.0%) pooled specificity 93.0% (95% CI 91.0–94.0%) MRI pooled sensitivity 77.0% (95% CI 65.0–87.0%) pooled specificity 92.0% (95% CI 88.0–95.0%)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? 8 studies out of 17 with consecutive enrollment

		Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear (12 studies out of 17 with retrospective or unclear design) Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Yes in 58.8% of studies Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	High risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? Unclear the clinical phase of patients (staging or suspected

		recurrence or both?); unclear the histological type of cancer; all studies consider bone metastasis only CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

8 PS - M staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Kruger 2011	FDG-PET/CT	104	patients at initial staging with non-small cell lung cancer (82) or small cell lung cancer (22), without suspected brain metastasis	27.3	97.6

PS - Kruger 2011

Clinical features and settings	non-small-cell lung cancer and small cell lung cancer; Country: Germany
Participants	104 patients (77 males; median age 65 years, range 44–87) at initial staging with non-small cell lung cancer (82) or small cell lung cancer (22), without suspected brain metastasis. 52 patients have stage IV disease. Prevalence of brain metastasis: 22 out of 104 patients
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	Brain metastasis Reference standard: MRI
Index and comparator tests	FDG-PET/CT (80 patients with contrast enhancement); comparator: none
Follow-up	not applicable
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Yes: consecutive Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Visual interpretation

		Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Low risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Yes Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? No: 17 patients excluded Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	High risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? 79% of participants with non-small cell lung cancer and 21% of participants with small cell lung cancer CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

9 PS - Any staging

Author, year	Design	Number of participants	Population	Follow up	Outcome	FDG-PET/CT arm (% of participants)	no FDG-PET/CT arm (% of participants)	p
Fischer 2009	open RCT	189 (98 FDG-PET/CT group; 91 conventional staging group)	suspected potentially resectable NSCLC	12 months	Operable	60 (61)	73 (80)	0.004
					Reason for nonoperability	11 (29)	0	0.05
					Stage IV disease	12 (31)	6 (33)	0.29
					Stage IIIB disease	14 (37)	12 (67)	0.15
					Stage IIIA (N2) disease	1 (3)	0	
					Other	21 (35)	38 (52)	
					Futile thoracotomy	5 (24)	4 (11)	
					Reason that thoracotomy was considered futile	0	8 (3)	
					Exploratory thoracotomy	3 (14)	0	
					Benign lung lesion	4 (19)	8 (21)	
					Stage IV disease	5 (24)	6 (16)	
					Stage IIIB disease	3 (14)	13 (34)	
					Stage IIIA (N2) disease	1 (5)	4 (11)	
					Recurrence within 12 mo	31 mo	49 mo	
					Death within 12 mo	60 (61)	46 (51)	
					median survival			
					death			

Maziak 2009	open RCT	337 (170 FDG-PET/CT group; 167 conventional staging group)	potentially resectable NSCLC	22 months	correctly upstaged incorrectly upstaged incorrectly understaged death	23 (13.8) 8 (4.8) 25 (14.9%) 52	11 (6.8) 1 (0.6) 48 (29.6%) 57	difference 7.0% (95% CI 0.3-13.7) P 0.046 difference 4.2% (95% CI, 0.5-8.6) P 0.037 difference, 14.7% (95% CI 5.7-23.4) P 0.002 HR, 0.88 (95% CI 0.61–1.29)
Fontaine 2011	cohort study	1999 (934 FDG-PET/CT group; 1065 no-FDG-PET/CT group)	patients undergoing resections for NSCLC	18 months	Overall survival* Stage Ia Stage Ib Stage II Stage III	61% (0.02 SE) 41% (0.09 SE)	53% (0.03 SE) 20% (0.05 SE)	0.04 n.s. n.s. n.s 0.03

PS - Fischer 2009

Clinical features and settings	non-small cell lung cancer; Country: Denmark
Participants	189 patients newly diagnosed or highly suspected for non-small cell lung-cancer, considered to have operable disease after conventional-staging procedures (i.e., medical history, physical examination, blood test, contrast-enhanced CT scan of the chest and upper abdomen, and bronchoscopy). After conventional staging, eligible patients were randomly assigned in a 1:1 ratio to PET–CT and conventional staging, followed by further invasive diagnostic procedures such as mediastinoscopy and endoscopic or endobronchial ultrasonography (the PET–CT group = 98 participants), or to conventional staging and invasive diagnostic procedures alone (the conventional-staging group = 91 participants). Eleven patients in the PET–CT group did not undergo PET–CT because of an unacceptably long waiting time for a scan or technical problems with the PET–CT equipment. One patient underwent PET–CT but declined all further staging procedures and surgery. Mediastinoscopy was performed in 89 patients in the

	PET–CT group (91%) and 88 in the conventional-staging group (97%) (P = 0.33). PET-CT group Conventional-staging group N N Age (yr) mean 63 64 range 42–80 38–80 Male sex (no.) 53 49 Female sex (no.) 45 42 ECOG performance status (no.) 0–1 93 86 2 1 1 Not available 4 4 Tumor Size on CT (mm) Mean 46.5 43.6 Range 10.0–110.0 15.0–130.0 TNM stage based on CT of thorax and abdomen (no.) IA 13 9 IB 17 13 IIA 0 0 IIB 5 7 IIIA 26 28 IIIB 32 32 IV 5 2 Mediastinoscopy (no.) Total 89 88 Stage N2 to N3 disease 9 12 Histologic features at operation (no) Squamous-cell carcinoma 22 22 Adenocarcinoma 30 29 Large-cell carcinoma 4 12 Bronchoalveolar carcinoma 0 1 NSCLC with no further specification 5 4 Other 2 2 Benign lung lesion 0 3
Study design	Open randomised controlled trial. Randomization was performed centrally with the use of a permuted-block design, stratified according to sex and recruiting center. Before a decision to operate was made, a consensus on the TNM stage was reached by a pulmonologist and a thoracic surgeon on the basis of all available information (clinical data, initial CT scanning, PET–CT imaging, bronchoscopy, mediastinoscopy, and if available, endoscopic ultrasonography with fine-needle aspiration or endobronchial ultrasonography). Mediastinoscopy and endoscopic or endobronchial ultrasonography served as the standard for

	preoperative assessment of mediastinal lymph nodes. All patients with stage I to stage IIB NSCLC were offered surgery. Patients with involvement of mediastinal lymph nodes or distant metastases (stage IIIA [N2] to stage IV) were considered to have inoperable disease and were offered chemotherapy with or without radiotherapy. Positive findings on PET–CT were further evaluated by biopsy or other imaging techniques (ultrasonography, radiography, or magnetic resonance imaging) at the discretion of the referring clinician. The study was closed after the inclusion of only 189 patients (expected 215 assigned to each group) because of slow accrual.
Target condition and reference standard(s)	N staging Clinical outcomes: the primary end point of the study was the frequency of futile thoracotomies (a benign lung lesion, pathologically proven mediastinal lymph-node involvement (stage IIIA [N2]), stage IIIB or IV disease, inoperable T3 or T4 disease, or recurrent disease or death from any cause within 1 year after randomization); median survival; death Reference standard: pathologist after thoracotomy served as the reference (N final). For patients in whom thoracotomy was not performed, N stage assigned by mediastinoscopy, EUS-FNA or EBUS-TBNA
Index and comparator tests	FDG-PET/CT (and conventional diagnostic tools for staging: clinical data, initial CT scanning, bronchoscopy); comparator: staging without FDG-PET/CT (clinical data, initial CT scanning, bronchoscopy)
Follow-up	12 months
Notes	No financial support was received from companies that make PET–CT scanners

Risk of bias table (Impact on clinical outcomes)

Bias	**Authors' judgement**
Random sequence generation (selection bias)	Low risk
Allocation concealment (selection bias)	Low risk
Blinding of participants and personnel (performance bias)	High risk
Blinding of outcome assessment (detection bias)	High risk
Incomplete outcome data (attrition bias)	Low risk
Selective reporting (reporting bias)	Low risk
Other bias	Low risk

PS - Maziak 2009

Clinical features and settings	non-small cell lung cancer; Country: Canada
Participants	337 patients diagnosed for non-small cell lung-cancer, considered to have operable disease (I, II, or IIIA disease) after staging procedures with chest CT. Eligible patients were randomly assigned in a 1:1 ratio to PET–CT and conventional staging, followed by further invasive diagnostic procedures such as mediastinoscopy and endoscopic or endobronchial ultrasonography (the PET–CT group = 170 participants), or to conventional staging and invasive diagnostic procedures alone (the conventional-staging group = 167 participants). Patients assigned to conventional staging underwent CT of the liver and adrenals (unless they were adequately visualized to rule out intra-abdominal metastases on the CT before randomization) and a whole-body bone scan. Patients in both groups underwent brain CT with contrast or brain magnetic resonance imaging with gadolinium. In patients whose imaging was negative for mediastinal disease, diagnostic confirmation by cervical mediastinoscopy was preferred. All patients, however, required detailed lymph node sampling at thoracotomy. PET-CT group Conventional-staging group N N Mean age (minimum, maximum), y 67 (41, 87) 66 (38, 88) Sex, n (%) Female 87 (51) 84 (50) Male 83 (49) 83 (50) Smoking status, n (%) Never 12 (7) 16 (10) Ex-smoker 110 (65) 106 (63) Current smoker 48 (28) 45 (27) ECOG performance status, n (%) 0 100 (59) 102 (61) 1 63 (37) 58 (35) 2 7 (4) 7 (4) Mean size of primary tumor (minimum, maximum), cm 3.2 (0.8, 8.7) 3.2 (0.9, 8.5) Clinical disease stage, n (%) IA 83 (49) 75 (45) IB 50 (29) 54 (32) IIA 6 (4) 2 (1) IIB 13 (8) 20 (12) IIIA 18 (10) 16 (10)
Study design	Open randomised controlled trial. An independent statistician created a computer generated randomization list, stratified by clinical stage (I or II vs. IIIA) and treatment center. A binder, which contained separate allocation sequences for each stratum, was kept in a locked drawer in the trial coordinator's office; access to this binder was limited to the coordinator and the data management assistant who provided the allocation. Within the binder, the allocation sequences were not

	concealed. In patients whose imaging was negative for mediastinal disease, diagnostic confirmation by cervical mediastinoscopy was preferred. All patients, however, required detailed lymph node sampling at thoracotomy. Patients with stage I, II, or IIIA disease underwent thoracotomy with resection of the primary lung lesion. Mediastinal node sampling appropriate for the lobe to be resected was performed at thoracotomy regardless of whether cervical mediastinoscopy had been performed. Patients underwent lung resection by open posterolateral thoracotomy or video-assisted thoracotomy. All patients had either lobectomy of the involved lobe or pneumonectomy, where appropriate. Postoperatively, patients could receive stage-appropriate adjuvant therapy (chemotherapy, radiotherapy, or a combination of these methods).
Target condition and reference standard(s)	The primary outcome was correct upstaging of cancer (true-positive results) where the imaging strategy identified a patient as having metastatic disease (stage IV) or locally advanced lung cancer (stage IIIB), thereby avoiding stage inappropriate surgery. Other outcomes included incorrect upstaging (false positive results) and incorrect understaging (false-negative results). The criteria used to define the latter outcome were pathologic stage IIIA or IIIB disease on mediastinoscopy or on lymph node sampling at thoracotomy or local recurrence or development of distant metastases within 1 year of thoracotomy (stage IV disease). Reference standard: in patients whose imaging was negative for mediastinal disease, diagnostic confirmation by cervical mediastinoscopy was preferred. All patients, however, required detailed lymph node sampling at thoracotomy. Patients with stage I, II, or IIIA disease underwent thoracotomy with resection of the primary lung lesion. Mediastinal node sampling appropriate for the lobe to be resected was performed at thoracotomy regardless of whether cervical mediastinoscopy had been performed. Upstaging of cancer was considered correct if the recommended further testing (biopsy or other diagnostic imaging modalities) confirmed it. Histopathologic confirmation during work-up of test abnormalities was not always required to label someone as correctly upstaged.
Index and comparator tests	PET-CT plus cranial imaging; comparator: conventional staging (abdominal CT, including the liver and adrenals, and bone scan) plus cranial imaging
Follow-up	The median duration of follow-up was 21.8 months (minimum, 0.1 month; maximum, 46.0 months) in the PET-CT group and 22.5 months (minimum, 0.2 month; maximum, 38.3 months) in the conventional staging group.
Notes	The trial was funded by the Ontario Ministry of Health and Long-Term Care, Canadian Institutes of Health Research, and Cancer Care Ontario

Assessment of methodological quality table (diagnostic accuracy)

Item	Authors' judgement	Support for judgement

1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Visual interpretation Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	High risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? No Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? No: 8 patients excluded Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of	Low risk	Describe included patients (prior testing, presentation, intended

applicability)		use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

Risk of bias (impact on clinical outcomes)

Bias	**Authors' judgement**
Random sequence generation (selection bias)	Low risk
Allocation concealment (selection bias)	Low risk
Blinding of participants and personnel (performance bias)	High risk
Blinding of outcome assessment (detection bias)	High risk
Incomplete outcome data (attrition bias)	Low risk
Selective reporting (reporting bias)	Low risk
Other bias	Low risk

PS - Fontaine 2011

Clinical features and settings	non-small-cell lung cancer; Country: UK
Participants	From a prospective validated thoracic surgery database, 1999 patients underwent a lung resection for proven or suspected non-small-cell lung cancer. No patients underwent resection with a preoperative stage N2, or received neo-adjuvant therapy to downstage. Staging was defined as pathological staging to eliminate bias by ‘better’ pre operative staging due to multislice computed tomography (CT) and PET/CT scanning. Mediastinsocopy was used in all patients who had mediastinal lymph nodes enlarged by CT criteria, or who had undergone PET scanning and were thought to have positive N2 nodes. Patients who had a PET/CT scan pre operatively (N = 934), were compared with patients who had not undergone PET/CT scanning (N = 1065) prior to surgical

	resection. PET/CT group PET/CT group N N P Preoperative characteristics Female (%) 488 (45.8) 463 (49.6) 0.09 Age at operation 68 (60—74) 69 (62—75) 0.0005 Smoking status (%) Current 353 (33.4) 284 (31.6) Ex 641 (60.7) 562 (62.5) Non 62 (5.9) 53 (5.9) 0.68 Pack years 40 (23—52) 40 (25—50) 0.66 Histology (%) 0.004 Adenoca 480 (45.1) 457 (48.9) Squamous 450 (42.3) 400 (42.8) Others 135 (12.7) 77 (8.2) Cancer stage (%) Ia 295 (27.7) 326 (34.9) Ib 408 (38.3) 344 (36.8) IIa 39 (3.7) 32 (3.4) IIb 173 (16.2) 132 (14.1) IIIa 102 (9.6) 70 (7.5) IIIb 37 (3.5) 23 (2.5) IV 11 (1.0) 7 (0.8) 0.02 Residual disease (%) 52 (4.9) 34 (3.6) 0.17
Study design	cohort study (thoracic surgery database with prospective enrollment of patients as data source)
Target condition and reference standard(s)	Outcomes: overall survival Reference standard: not applicable
Index and comparator tests	FDG-PET/CT for staging; comparator: staging without FDG-PET/CT
Follow-up	1.5 years (median) in the PET/CT group and 3.7 years in the non-PET/CT group
Notes	A Cox regression model to predict survival for all stages and then for each stage (I, II and III) was developed individually. Significant predictors of survival (and those approaching significance ($p < 0.1$)) were used to propensity match patients from the PET/CT and no-PET/CT groups.

Assessment of methodological quality table

Bias	Authors' judgement
Random sequence generation (selection bias)	High risk
Allocation concealment (selection bias)	High risk
Blinding of participants and personnel (performance bias)	Unclear risk
Blinding of outcome assessment (detection bias)	Unclear risk
Incomplete outcome data (attrition bias)	Unclear risk
Selective reporting (reporting bias)	Low risk
Other bias	Low risk

Characteristics of excluded studies

Borekci 2011

Reason for exclusion	per-lesion analysis

Callaway 2010

Reason for exclusion	abstract at congress

Cao 2011

Reason for exclusion	cost-effectiveness study

Ceylan 2012

Reason for exclusion	retrospective study

Chen 2010

Reason for exclusion	per-lesion analysis

Cho 2011

Reason for exclusion	retrospective study

De Wever 2010

Reason for exclusion	retrospective study

Detterbeck 2010

Reason for exclusion	letter

Eloubeidi 2011

Reason for exclusion	abstract at congress

Eloubeidi 2011a

Reason for exclusion	abstract at congress

Fischer 2011

Reason for exclusion	letter

Geraldson 2012

Reason for exclusion	retrospective study

Gomez-Caro 2010

Reason for exclusion	sample of FDG-PET/CT negative patients

Gulenchyn 2010

Reason for exclusion	abstract at congress

Gunluoglu 2010

Reason for exclusion	abstract at congress

Haak-Siepel 2011

Reason for exclusion	abstract at congress

Harders 2011

Reason for exclusion	abstract at congress

Herbrik 2011

Reason for exclusion	retrospective study

Herder 2006

Reason for exclusion	not FDG-PET/CT

Heusner 2011

Reason for exclusion	mixed population of cancers (non-small cell lung cancer and melanoma)

Hsu 2011

Reason for exclusion	abstract at congress

Hu 2011

Reason for exclusion	unclear design (prospective or retrospective)

Iskender 2011

Reason for exclusion	per-lesion analysis

Jayaram 2010

Reason for exclusion	abstract at congress

Jeon 2010

Reason for exclusion	retrospective study

Jung 2012

Reason for exclusion	retrospective study

Kasai 2010

Reason for exclusion	unclear design (prospective or retrospective)

Kim 2011

Reason for exclusion	abstract at congress

Kim 2012

Reason for exclusion	per-node analysis

Kubota 2011

Reason for exclusion	not FDG-PET/CT

Kubota 2011a

Reason for exclusion	not staging

Kuo 2011

Reason for exclusion	not FDG-PET/CT

Langer 2010

Reason for exclusion	economic study

Lee 2011

Reason for exclusion	retrospective study

Lee 2012

Reason for exclusion	retrospective study

Li 2012

Reason for exclusion	retrospective study

Lin 2010

Reason for exclusion	retrospective study

Mac Manus 2010

Reason for exclusion	narrative review

Metin 2011

Reason for exclusion	retrospective study

Moralejo 2010

Reason for exclusion	sensitivity and specificity estimates not reported

Nambu 2010

Reason for exclusion	retrospective study

Navani 2010

Reason for exclusion	editorial

Navani 2010a

Reason for exclusion	sensitivity and specificity estimates not reported

Okereke 2011

Reason for exclusion	retrospective study

Ose 2012

Reason for exclusion	retrospective study

Ozcan 2011

Reason for exclusion	mixed population of several cancers

Ozhan 2011

Reason for exclusion	abstract at congress

Paesmans 2010

Reason for exclusion	prognostic study

Pauls 2012

Reason for exclusion	FDG-PET as reference standard

Peng 2011

Reason for exclusion	choline PET

Pepek 2011

Reason for exclusion	abstract at congress

Portilla-Quattrociocchi 2011

Reason for exclusion	retrospective study

Pulvirenti 2010

Reason for exclusion	retrospective study

Ruben 2011

Reason for exclusion	abstract at congress

Sanchez 2011

Reason for exclusion	retrospective study

Saw 2011

Reason for exclusion	abstract at congress

Selvaraj 2011

Reason for exclusion	

Sogaard 2011

Reason for exclusion	abstract at congress

Song 2011

Reason for exclusion	retrospective study

Spaggiari 2005

Reason for exclusion	retrospective study

Spiegler 2011

Reason for exclusion	full-text not found

Tasci 2010

Reason for exclusion	per-lesion analysis

Tupayachi 2010

Reason for exclusion	abstract at congress

Tupayachi 2010a

Reason for exclusion	abstract at congress

Ung 2011

Reason for exclusion	abstract at congress

Usuda 2011

Reason for exclusion	per-lesion analysis

van Tinteren 2002

Reason for exclusion	not FDG-PET/CT

van't Westeinde 2011

Reason for exclusion	study on primary diagnosis

Vaz 2012

Reason for exclusion	retrospective study

Ventura 2010

Reason for exclusion	retrospective study

Viney 2004

Reason for exclusion	cost-effectiveness study

Wang 2012

Reason for exclusion	systematic review only on negative predictive value

Yang 2010

Reason for exclusion	per-node analysis

Yi 2011

Reason for exclusion	retrospective study

Zsiray 2011

Reason for exclusion	abstract at congress

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”: ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. Lung NEAR Cancer*: ti,ab,kw
10. Pulmonary nodule*: ti,ab,kw
11. “Lung neoplasms”[Mesh explodes all trees]
12. 9/11 OR
13. 8 AND 12

Publication date: January 2010 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18”[Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg pet”[All Fields]
17. “Positron-Emission Tomography”[Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. “Lung Neoplasms”[Mesh:noexp]
26. “Bronchial Neoplasms”[Mesh])
27. “Multiple Pulmonary Nodules”[Mesh]
28. “Solitary Pulmonary Nodule”[Mesh]
29. “non-small cell lung cancer”[Title/Abstract]
30. “non-small cell lung carcinoma”[Title/Abstract]
31. “non-small cell lung carcinomas”[Title/Abstract]
32. “non-small lung cancers”[Title/Abstract]
33. “lung cancer”[Title/Abstract]
34. “pulmonary cancer”[Title/Abstract]
35. “pulmonary cancers”[Title/Abstract]
36. “lung cancers”[Title/Abstract]
37. “bronchogenic carcinoma”[Title/Abstract]
38. “bronchogenic carcinomas”[Title/Abstract]
39. “bronchial carcinoma”[Title/Abstract]
40. “bronchial carcinomas”[Title/Abstract]
41. “small cell lung cancer”[Title/Abstract]
42. “small cell lung cancers”[Title/Abstract]

43. “multiple pulmonary nodules”[Title/Abstract]
44. “solitary pulmonary nodule”[Title/Abstract]
45. “solitary pulmonary nodules”[Title/Abstract]
46. “solitary pulmonary tumor”[Title/Abstract]
47. “solitary pulmonary tumors”[Title/Abstract]
48. “pulmonary coin lesion”[Title/Abstract]
49. “pulmonary coin lesions”[Title/Abstract]
50. 25/49 OR
51. 24 AND 50

Limit: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2010 - March 2012

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “fluorodeoxyglucose f 18”/exp
3. “fluorodeoxyglucose f 18”/syn
4. “computer assisted emission tomography”/exp
5. “computer assisted emission tomography” OR
6. pet
7. “pet scans”
8. “pet scanner”
9. “pet scan”
10. “pet/ct scan”
11. “pet/ct scans”
12. “pet/ct”
13. “positron emission tomography/computed tomography”
14. pet NEAR/4 scan*
15. pet NEAR/4 ct
16. 1/15 OR
17. “lung cancer”/ de,syn, keyword
18. “lung metastasis”/ de,syn, keyword
19. “lung sarcoma”/ de,syn, keyword
20. “lung nodule”/ de,syn, keyword
21. “lung metastasis”/ de,syn, keyword
22. “lung sarcoma”/ de,syn, keyword
23. “lung nodule”/ de,syn, keyword
24. “lung carcinoma”/ de,syn, keyword
25. “lung carcinoma”/ de,syn, keyword
26. lung adenocarcinoma/ de,syn, keyword
27. lung alveolus cell carcinoma/ de,syn, keyword
28. lung non-small cell cancer/ de,syn, keyword
29. lung small cell cancer/ de,syn, keyword
30. lung squamous cell carcinoma/ de,syn, keyword
31. “lung nodule”: ab,ti
32. “pulmonary nodule”: ab,ti
33. “lung cancer”: ab,ti
34. “pulmonary cancer”: ab,ti
35. “lung metastastis”: ab,ti
36. “bronchopulmonary metastasis”: ab,ti
37. “bronchus metastasis”: ab,ti
38. “lung near/3 sarcoma”: ab,ti
39. “lung alveolus sarcoma”: ab,ti
40. “malignant lung sarcoma”: ab,ti
41. “pulmonary sarcoma”: ab,ti
42. “bronchial carcinoma”: ab,ti
43. “lung sarcoma”: ab,ti
44. “bronchopulmonary carcinoma”: ab,ti
45. “bronchus carcinoma”: ab,ti
46. “lung carcinoma”: ab,ti
47. “pulmonary adenocarcinoma”: ab,ti
48. “alveobronchial carcinoma”: ab,ti

49. “lobular carcinoma”: ab,ti
50. “lung cavitary carcinoma”: ab,ti
51. “peribronchial carcinoma”: ab,ti
52. “lung alveolus cell carcinoma”: ab,ti
53. “alveolar carcinoma”: ab,ti
54. “bronchioalveolar lung carcinoma”: ab,ti
55. “bronchoalveolar carcinoma”: ab,ti
56. “bronchoalveolar cancer”: ab,ti
57. “alveolar cell cancer”: ab,ti
58. “alveolar cell carcinoma”: ab,ti
59. “lung alveolus cell cancer”: ab,ti
60. “pulmonary alveolar cell cancer”: ab,ti
61. “lung non-small cell cancer”: ab,ti
62. “non-small-cell lung cancer”: ab,ti
63. “lung small cell cancer”: ab,ti
64. “small cell lung carcinoma”: ab,ti
65. “small cell lung cancer”: ab,ti
66. “lung squamous cell carcinoma”: ab,ti
67. “lung epidermoid cancer”: ab,ti
68. “lung squamous cell cancer”: ab,ti
69. 17/68 OR
70. 16 AND 69
71. 70 AND (“article” OR “review”/it OR “short survey”)

Limit: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2010 - March 2012

APPENDIX 4

FDG-PET/CT for staging of small cell lung cancer

Characteristics of included studies

HTA report - ASSR 2012 SCLC

Document ID	ASSR-RER 2012 - Lung cancer
Objectives	to define criteria for appropriate use of FDG-PET for patients with small cell lung cancer
Methods	A panel of experts working in Health Trusts and Teaching Hospitals of Emilia-Romagna was convened to discuss and agree on the methodology for a research programme aimed at defining the criteria for appropriate use of PET in small cell lung cancer. On the basis of the clinical pathway of patients with small cell lung cancer the panel examined and assessed the role of FDG-PET for 4 clinical indications (staging of patients with primary small cell lung cancer, target volume definition of radiation treatment with curative intent in patients treated for lung cancer, during-treatment evaluation of response to systemic therapy in patients treated for small cell lung cancer, end of treatment evaluation of response to systemic therapy in patients treated for small cell lung cancer). The following databases were searched for the period between January 2006 and September 2010: Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE - The Cochrane Library); Health Technology Assessment Database; Cochrane Central Register of Controlled Trials; National Library of Medicine's Medline database (PubMed); Elsevier's Embase. Language restrictions: English, Italian, French and Spanish. Selection criteria Type of studies: systematic reviews, RCTs, CCTs, cross-sectional diagnostic studies, prospective or retrospective cohort studies, case series of at least 10 patients Participants: patients with small cell lung cancer Intervention: FDG-PET or CT/PET Reference standard: histology or clinical follow up Comparator: any other imaging technique Outcomes> sensitivity, specificity, LR, metabolic/tumor response, time to recurrence, local, local-regional and distant recurrence, disease free survival, disease survival, overall survival Assessment of methodological quality of studies The following criteria have been used for the quality assessment of different study designs. Systematic reviews: criteria drawn from the AMSTAR checklist Diagnostic cross sectional studies: criteria drawn from the QUADAS checklist Randomized controlled trials: criteria suggested by the Cochrane Handbook

	Case control studies and cohort studies: criteria drawn from the New Castle-Ottawa checklist Level of evidence for estimates of diagnostic accuracy were assigned according to GRADE categorization of the quality of evidence Each member of the panel voted the level of appropriateness for each clinical question. Two rounds of votes were requested for the judgment of appropriateness and results were analysed using the RAND/UCLA Appropriateness Method. The use of FDG-PET for a specific clinical indication was judged was judged as *appropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 7-9 score region as *inappropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 1-3 score region. Finally the use of FDG-PET was judged as *uncertain* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 4-6 score region or when no agreement was reached after the second round of voting. Clinical indications for which the panel does not reach an agreement on level of appropriateness after two rounds of voting also fall in the *uncertain* category.
Conclusions	STAGING OF PATIENTS WITH SMALL CELL LUNG CANCER - SCLC - UNCERTAIN The available data on FDG-PET accuracy in discriminating limited from extended SCLC are sparse and the level of evidence was considered very low. The limited difference in gain offered by the therapeutic options available led the panel to give low scores for clinical outcomes: consequences for true and false positive treated with just chemotherapy were voted not important (median score 3), while consequences for true and false negative receiving combined chemo/radiotherapy were voted important (median score 4). Both voting rounds on appropriateness registered a disagreement among panelists with ratings falling in both the inappropriate and uncertain regions. The use of FDG-PET in staging SCLC resulted therefore uncertain due to disagreement.
Notes	Meta-analysis of diagnostic accuracy estimates was not performed

SR - Qu 2011 - M staging

Disease	non-small cell lung cancer and other type of lung cancer
Index test	FDG-PET or FDG-PET/CT
Comparators	bone scintigraphy
Reference standard	histopathology alone, histopathology plus clinical follow-up or clinical follow-up alone
Target	diagnostic accuracy for bone metastasis
Studies included	diagnostic accuracy primary studies with prospective or retrospective design (not clear if studies on staging or recurrence or both)
Years covered by the search	up to August 2010
Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE, EMBASE, Cochrane Library)
Characteristics of included studies clearly reported in tables	Partially (not reported neither the clinical phase - staging / recurrence- nor the histological type of cancer)
Methodological quality of primary studies assessed; criteria reported	Yes (QUADAS tool)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	17 primary studies (11 including only NSCLC patients; 5 including both NSCLC and SCLC patients; 1 including only SCLC patients) 7 studies of FDG-PET/CT using patient as unit of analysis (2 with small cell lung cancer patients); 2 studies of FDG-PET/CT using lesion as unit of analysis; 5 studies of FDG-PET using patient as unit of analysis; 5 studies of FDG-PET/CT using lesion as unit of analysis; 3 studies of MRI using patient as unit of analysis; 3 studies of MRI using lesion as unit of analysis; 12 studies of bone scintigraphy using patient as unit of analysis (4 with small cell lung cancer patients); 4 studies of bone scintigraphy using lesion as unit of analysis
Design of included studies	Cross sectional diagnostic accuracy studies (5 studies out of 17 with prospective design; the other studies with retrospective or unclear design)

N. of included patients	2940; Considering studies using patient as unit of analysis: studies assessing FDG-PET/CT 1855 patients (small cell lung cancer patients: 211), studies assessing bone scintigraphy 2285 patients (small cell lung cancer patients: 645)
Diagnostic accuracy results (with heterogeneity)	Per-patient based analysis FDG-PET/CT (small cell lung cancer) pooled sensitivity 90.0% (95% CI 76.0–97.0%) pooled specificity 95.0% (95% CI 90.0–98.0%) bone scintigraphy (small cell lung cancer) pooled sensitivity 88.0% (95% CI 81.0–93.0%) pooled specificity 74.0% (95% CI 70.0–77.0%)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? 8 studies out of 17 with consecutive enrollment Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear (12 studies out of 17 with retrospective or unclear design) Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of

		the results of the index test? Yes in 58.8% of studies Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	High risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? Unclear the clinical phase of patients (staging or suspected recurrence or both?); unclear the histological type of cancer; all studies consider bone metastasis only CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - M staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Kruger 2011	FDG-PET/CT	104	patients at initial staging with non-small cell lung cancer (82) or small cell lung cancer (22), without suspected brain metastasis	27.3	97.6

PS - Kruger 2011

Clinical features and settings	non-small-cell lung cancer and small cell lung cancer; Country: Germany
Participants	104 patients (77 males; median age 65 years, range 44–87) at initial staging with non-small cell lung cancer (82) or small cell lung cancer (22), without suspected brain metastasis. 52 patients have stage IV disease. Prevalence of brain metastasis: 22 out of 104 patients
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	Brain metastasis Reference standard: MRI
Index and comparator tests	FDG-PET/CT (80 patients with contrast enhancement); comparator: none
Follow-up	not applicable
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Yes: consecutive Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Visual interpretation Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Low risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Yes Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? No: 17 patients excluded Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	High risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? 79% of participants with non-small cell lung cancer and 21% of participants with small cell lung cancer CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question?

applicability)		CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

Characteristics of excluded studies

Borekci 2011

Reason for exclusion	per-lesion analysis

Callaway 2010

Reason for exclusion	abstract at congress

Cao 2011

Reason for exclusion	cost-effectiveness study

Ceylan 2012

Reason for exclusion	retrospective study

Chang 2012

Reason for exclusion	non-small cell lung cancer

Chen 2010

Reason for exclusion	per-lesion analysis

Cho 2011

Reason for exclusion	retrospective study

Darling 2011

Reason for exclusion	non-small cell lung cancer

De Wever 2010

Reason for exclusion	retrospective study

Detterbeck 2010

Reason for exclusion	letter

Eloubeidi 2011

Reason for exclusion	abstract at congress

Eloubeidi 2011a

Reason for exclusion	abstract at congress

Fischer 2009

Reason for exclusion	non-small cell lung cancer

Fischer 2011

Reason for exclusion	letter

Fontaine 2011

Reason for exclusion	non-small cell lung cancer

Geraldson 2012

Reason for exclusion	retrospective study

Gomez-Caro 2010

Reason for exclusion	sample of FDG-PET/CT negative patients

Gulenchyn 2010

Reason for exclusion	abstract at congress

Gunluoglu 2010

Reason for exclusion	abstract at congress

Gunluoglu 2011

Reason for exclusion	non-small cell lung cancer

Haak-Siepel 2011

Reason for exclusion	abstract at congress

Harders 2011

Reason for exclusion	abstract at congress

Herbrik 2011

Reason for exclusion	retrospective study

Herder 2006

Reason for exclusion	not FDG-PET/CT

Heusner 2011

Reason for exclusion	mixed population (non-small cell lung cancer and melanoma)

Hsu 2011

Reason for exclusion	abstract at congress

Hu 2011

Reason for exclusion	unclear design (prospective or retrospective)

Iskender 2011

Reason for exclusion	per-lesion analysis

Jayaram 2010

Reason for exclusion	abstract at congress

Jeon 2010

Reason for exclusion	retrospective study

Jung 2012

Reason for exclusion	retrospective study

Kasai 2010

Reason for exclusion	unclear design (prospective or retrospective)

Kim 2011

Reason for exclusion	abstract at congress

Kim 2012

Reason for exclusion	per-node analysis

Kubota 2011

Reason for exclusion	not FDG-PET/CT

Kubota 2011a

Reason for exclusion	not staging

Kuo 2011

Reason for exclusion	not FDG-PET/CT

Langer 2010

Reason for exclusion	economic study

Lee 2011

Reason for exclusion	retrospective study

Lee 2012

Reason for exclusion	retrospective study

Li 2011

Reason for exclusion	non-small cell lung cancer

Li 2012

Reason for exclusion	retrospective study

Lin 2010

Reason for exclusion	retrospective study

Liu 2011

Reason for exclusion	non-small cell lung cancer

Lv 2011

Reason for exclusion	non-small cell lung cancer

Mac Manus 2010

Reason for exclusion	narrative review

Maziak 2009

Reason for exclusion	non-small cell lung cancer

Metin 2011

Reason for exclusion	retrospective study

Moralejo 2010

Reason for exclusion	non-small cell lung cancer

Nambu 2010

Reason for exclusion	retrospective study

Navani 2010

Reason for exclusion	editorial

Navani 2010a

Reason for exclusion	non-small cell lung cancer

Ohnishi 2011

Reason for exclusion	non-small cell lung cancer

Ohno 2011

Reason for exclusion	non-small cell lung cancer

Okereke 2011

Reason for exclusion	retrospective study

Ose 2012

Reason for exclusion	retrospective study

Ozcan 2011

Reason for exclusion	mixed population of several cancers

Ozhan 2011

Reason for exclusion	abstract at congress

Paesmans 2010

Reason for exclusion	prognostic study

Pauls 2012

Reason for exclusion	FDG-PET as reference standard

Peng 2011

Reason for exclusion	choline PET

Pepek 2011

Reason for exclusion	abstract at congress

Portilla-Quattrociocchi 2011

Reason for exclusion	retrospective study

Pulvirenti 2010

Reason for exclusion	retrospective study

Ruben 2011

Reason for exclusion	abstract at congress

Sanchez 2011

Reason for exclusion	retrospective study

Saw 2011

Reason for exclusion	abstract at congress

Selvaraj 2011

Reason for exclusion	abstract at congress

Sivrikoz 2011

Reason for exclusion	non-small cell lung cancer

Sogaard 2011

Reason for exclusion	cost-effectiveness study

Song 2011

Reason for exclusion	retrospective study

Spaggiari 2005

Reason for exclusion	retrospective study

Spiegler 2011

Reason for exclusion	full-text not found

Tasci 2010

Reason for exclusion	per-lesion analysis

Tupayachi 2010

Reason for exclusion	abstract at congress

Tupayachi 2010a

Reason for exclusion	abstract at congress

Ung 2011

Reason for exclusion	abstract at congress

Usuda 2011

Reason for exclusion	per-lesion analysis

van Tinteren 2002

Reason for exclusion	not FDG-PET/CT

van't Westeinde 2011

Reason for exclusion	study on primary diagnosis

Vaz 2012

Reason for exclusion	retrospective study

Ventura 2010

Reason for exclusion	retrospective study

Viney 2004

Reason for exclusion	cost-effectiveness study

Wang 2012

Reason for exclusion	non-small cell lung cancer

Yang 2010

Reason for exclusion	non-small cell lung cancer

Yi 2011

Reason for exclusion	retrospective study

Zhao 2011

Reason for exclusion	non-small cell lung cancer

Zsiray 2011

Reason for exclusion	abstract at congress

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”: ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. Lung NEAR Cancer*: ti,ab,kw
10. Pulmonary nodule*: ti,ab,kw
11. “Lung neoplasms”[Mesh explodes all trees]
12. 9/11 OR
13. 8 AND 12

Publication date: January 2010 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18”[Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg pet”[All Fields]
17. “Positron-Emission Tomography”[Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. “Lung Neoplasms”[Mesh:noexp]
26. “Bronchial Neoplasms”[Mesh])
27. “Multiple Pulmonary Nodules”[Mesh]
28. “Solitary Pulmonary Nodule”[Mesh]
29. “non-small cell lung cancer”[Title/Abstract]
30. “non-small cell lung carcinoma”[Title/Abstract]
31. “non-small cell lung carcinomas”[Title/Abstract]
32. “non-small lung cancers”[Title/Abstract]
33. “lung cancer”[Title/Abstract]
34. “pulmonary cancer”[Title/Abstract]
35. “pulmonary cancers”[Title/Abstract]
36. “lung cancers”[Title/Abstract]
37. “bronchogenic carcinoma”[Title/Abstract]
38. “bronchogenic carcinomas”[Title/Abstract]
39. “bronchial carcinoma”[Title/Abstract]
40. “bronchial carcinomas”[Title/Abstract]
41. “small cell lung cancer”[Title/Abstract]
42. “small cell lung cancers”[Title/Abstract]
43. “multiple pulmonary nodules”[Title/Abstract]

44. “solitary pulmonary nodule”[Title/Abstract]
45. “solitary pulmonary nodules”[Title/Abstract]
46. “solitary pulmonary tumor”[Title/Abstract]
47. “solitary pulmonary tumors”[Title/Abstract]
48. “pulmonary coin lesion”[Title/Abstract]
49. “pulmonary coin lesions”[Title/Abstract]
50. 25/49 OR
51. 24 AND 50

Limit: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2010 - March 2012

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “fluorodeoxyglucose f 18”/exp
3. “fluorodeoxyglucose f 18”/syn
4. “computer assisted emission tomography”/exp
5. “computer assisted emission tomography” OR
6. pet
7. “pet scans”
8. “pet scanner”
9. “pet scan”
10. “pet/ct scan”
11. “pet/ct scans”
12. “pet/ct”
13. “positron emission tomography/computed tomography”
14. pet NEAR/4 scan*
15. pet NEAR/4 ct
16. 1/15 OR
17. “lung cancer”/ de,syn, keyword
18. “lung metastasis”/ de,syn, keyword
19. “lung sarcoma”/ de,syn, keyword
20. “lung nodule”/ de,syn, keyword
21. “lung metastasis”/ de,syn, keyword
22. “lung sarcoma”/ de,syn, keyword
23. “lung nodule”/ de,syn, keyword
24. “lung carcinoma”/ de,syn, keyword
25. “lung carcinoma”/ de,syn, keyword
26. lung adenocarcinoma/ de,syn, keyword
27. lung alveolus cell carcinoma/ de,syn, keyword
28. lung non-small cell cancer/ de,syn, keyword
29. lung small cell cancer/ de,syn, keyword
30. lung squamous cell carcinoma/ de,syn, keyword
31. “lung nodule”: ab,ti
32. “pulmonary nodule”: ab,ti
33. “lung cancer”: ab,ti
34. “pulmonary cancer”: ab,ti
35. “lung metastastis”: ab,ti
36. “bronchopulmonary metastasis”: ab,ti
37. “bronchus metastasis”: ab,ti
38. “lung near/3 sarcoma”: ab,ti
39. “lung alveolus sarcoma”: ab,ti
40. “malignant lung sarcoma”: ab,ti
41. “pulmonary sarcoma”: ab,ti
42. “bronchial carcinoma”: ab,ti
43. “lung sarcoma”: ab,ti
44. “bronchopulmonary carcinoma”: ab,ti
45. “bronchus carcinoma”: ab,ti
46. “lung carcinoma”: ab,ti
47. “pulmonary adenocarcinoma”: ab,ti
48. “alveobronchial carcinoma”: ab,ti

49. “lobular carcinoma”: ab,ti
50. “lung cavitary carcinoma”: ab,ti
51. “peribronchial carcinoma”: ab,ti
52. “lung alveolus cell carcinoma”: ab,ti
53. “alveolar carcinoma”: ab,ti
54. “bronchioalveolar lung carcinoma”: ab,ti
55. “bronchoalveolar carcinoma”: ab,ti
56. “bronchoalveolar cancer”: ab,ti
57. “alveolar cell cancer”: ab,ti
58. “alveolar cell carcinoma”: ab,ti
59. “lung alveolus cell cancer”: ab,ti
60. “pulmonary alveolar cell cancer”: ab,ti
61. “lung non-small cell cancer”: ab,ti
62. “non-small-cell lung cancer”: ab,ti
63. “lung small cell cancer”: ab,ti
64. “small cell lung carcinoma”: ab,ti
65. “small cell lung cancer”: ab,ti
66. “lung squamous cell carcinoma”: ab,ti
67. “lung epidermoid cancer”: ab,ti
68. “lung squamous cell cancer”: ab,ti
69. 17/68 OR
70. 16 AND 69
71. 70 AND (“article” OR “review”/it OR “short survey”)

Limit: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2010 - March 2012

APPENDIX 5

FDG-PET/CT for staging of malignant pleural mesothelioma

Characteristics of included studies

HTA report - KCE 2009

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients;

	• Absence of adequate reference standard; • Absence of patient-based analysis; • Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	The evidence on the use of PET for mesothelioma is limited to one primary study and does not allow the formulation of firm conclusions (level 2).
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

PS - T staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Erasmus 2005	FDG-PET/CT	29	malignant pleural mesothelioma considered for extrapleural pneumonectomy	67.0	93.0
Pilling 2010	FDG-PET/CT	20	malignant pleural mesothelioma treated with extrapleural pneumonectomy	0	100
Sorensen 2008	FDG-PET/CT	24	malignant pleural mesothelioma considered for extrapleural pneumonectomy	75.0	100

PS - N staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Erasmus 2005	FDG-PET/CT	17	malignant pleural mesothelioma considered for extrapleural pneumonectomy	38.0	78.0
Pilling 2010	FDG-PET/CT	20	malignant pleural mesothelioma treated with extrapleural pneumonectomy	11.1	93.3
Sorensen 2008	FDG-PET/CT	30	malignant pleural mesothelioma considered for extrapleural pneumonectomy	50.0	75.0

PS - Any staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Erasmus 2005	FDG-PET/CT	29	malignant pleural mesothelioma considered for extrapleural pneumonectomy	85.7	73.3

PS - Erasmus 2005

Clinical features and settings	malignant pleural mesothelioma; Country: US
Participants	29 patients (26 men and 3 women; mean age 63 years [range, 44-77 years]) with malignant pleural mesothelioma considered for extrapleural pneumonectomy followed by intensity modulated radiation therapy (stage I to III = T1-3 N1-2 M0).
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	T staging (T4 disease vs other T disease) N staging (N2 disease vs other N disease) = only 17 patients for this target condition Any staging (stage I to III = considered for extrapleural pneumonectomy, versus stage IV = not considered for extrapleural pneumonectomy) Reference standard: histopathology and/or results of further radiologic evaluation or follow-up
Index and comparator tests	FDG-PET/CT; comparator: none
Follow-up	not reported
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Unclear if consecutive Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Visual interpretation Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Pilling 2010

Clinical features and settings	malignant pleural mesothelioma; Country: UK

Participants	20 patients (18 male, median age 62 years, range 52–68) who underwent extrapleural pneumonectomy as part of trimodality therapy for malignant pleural mesothelioma
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	T staging (T4 disease vs other T disease) N staging (N2 disease vs other N disease) Reference standard: surgical pathological status
Index and comparator tests	FDG-PET/CT; comparator: none
Follow-up	not reported
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Yes consecutive Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Visual interpretation Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

		Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	High risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? No: PET-CT scans were performed a median of 119 days (range 2–229) before the day of operation Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Sorensen 2008

Clinical features and settings	malignant pleural mesothelioma; Country: Denmark
Participants	42 patients (39 male, median age 61 years, range 30–70) with malignant pleural mesothelioma considered for extrapleural pneumonectomy.
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and	T staging (T4 disease vs other T disease) - analysis on 24 participants

reference standard(s)	N staging (N2 disease vs other N disease) - analysis on 30 participants Reference standard: final histological results obtained by mediastinoscopy and surgical/pathological results toghether
Index and comparator tests	FDG-PET/CT; comparator: none
Follow-up	not reported
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Yes consecutive Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING	High risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer

(risk of bias)		to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? No Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

Characteristics of excluded studies

Abe 2012

Reason for exclusion	retrospective design

Alvarez 2009

Reason for exclusion	only FDG-PET

Ambrosini 2005

Reason for exclusion	not a diagnostic design

Basu 2011

Reason for exclusion	narrative review

Dhalluin 2009

Reason for exclusion	document of recommendations

Kruger 2007

Reason for exclusion	not a diagnostic design

Lee 2009

Reason for exclusion	retrospective design; prognostic purpose

Lequaglie 2011

Reason for exclusion	abstract at congress

Plathow 2008

Reason for exclusion	lesion as unit of analysis

Sharif 2011

Reason for exclusion	narrative review

Wilcox 2009

Reason for exclusion	not a diagnostic design

Zahid 2011

Reason for exclusion	narrative review

Appendices

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18" [MeSH descriptor explode all trees]
3. "positron emission tomography":ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. "Fluorodeoxyglucose F18": ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. mesothelioma* OR mesothelial: all fields
10. Mesothelioma [MeSH descriptor explode all trees]
11. Neoplasms, Mesothelial [MeSH descriptor explode all trees]
12. 9/11 OR
13. 12 AND 8

Publication date: January 2009 - March 2012

2 MEDLINE search strategy

1. "Fluorodeoxyglucose F18"[Mesh]
2. "2-Fluoro-2-deoxyglucose" [All Fields]
3. "18F Fluorodeoxyglucose" [All Fields]
4. "F 18 Fluorodeoxyglucose" [All Fields]
5. Fludeoxyglucose* [All Fields]
6. "2 fluoro 2 deoxy d glucose"[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. "fluorine 18 fluorodeoxyglucose" [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. "fdg pet"[All Fields]
17. "Positron-Emission Tomography"[Mesh]
18. "positron emission tomography" [title/abstract]
19. pet [title/abstract]
20. "pet scan" [All Fields]
21. "pet scans" [All Fields]
22. "pet scanner" [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. "mesothelioma"[Title/Abstract]
26. Neoplasms, Mesothelial"[Mesh:NoExp]
27. "Mesothelioma"[Mesh]
28. 25/27 OR
29. 24 AND 28

Limit: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2009 - March 2012

3 EMBASE search strategy

1. 'positron emission tomography'/syn
2. 'fluorodeoxyglucose f 18'/exp
3. 'fluorodeoxyglucose f 18'/syn
4. 'computer assisted emission tomography'/exp

5. 'computer assisted emission tomography' OR
6. pet
7. 'pet scans'
8. 'pet scanner'
9. ‘pet scan'
10. 'pet/ct scan'
11. 'pet/ct scans'
12. 'pet/ct'
13. 'positron emission tomography/computed tomography'
14. pet NEAR/4 scan*
15. pet NEAR/4 ct
16. 1/23 OR
17. mesothelioma* OR mesothelial OR 'malignant mesothelioma' OR 'mesothelioma' OR 'pleura mesothelioma' [tw]
18. 'malignant mesothelioma'/exp OR 'mesothelioma'/exp OR 'pleura mesothelioma'/exp OR
19. 'malignant mesothelioma'/syn OR 'mesothelioma'/syn OR 'pleura mesothelioma'/syn
20. 17/19 OR
21. 20 AND 16

Limit: Humans embase only;
Publication type: Article, Article in press, Short review, Review
Languages: English, French, Italian, Spanish
Publication date: January 2009 - March 2012

APPENDIX 6

FDG-PET/CT for staging of breast cancer

Characteristics of included studies

HTA report - ASSR-RER 2011

Document ID	Ballini L, Vignatelli L, Negro A, Minozzi S, Maltoni S, Longo G. Criteria for appropriate use of FDG/PET in breast cancer. Dossier 207 - Agenzia sanitaria e sociale regionale, Regione Emilia-Romagna. 2011.
Objectives	to define criteria for appropriate use of FDG-PET for patients with breast cancer
Methods	A panel of experts working in Health Trusts and Teaching Hospitals of Emilia-Romagna was convened to discuss and agree on the methodology for a research programme aimed at defining the criteria for appropriate use of PET in breast cancer. On the basis of the clinical pathway of patients with breast cancer the panel examined and assessed the role of FDG-PET for 7 clinical indications (diagnosis of primary breast cancer, N staging of primary breast cancer, M staging of locally advanced breast cancer, evaluation of early response to neo-adjuvant therapy, evaluation of response to neo-adjuvant therapy at the end of treatment, follow up in patients with no suspicion of recurrence, diagnosis and staging of suspect distant recurrence). The following databases were searched for the period between January 2006 and July 2010: Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE - The Cochrane Library); Health Technology Assessment Database; Cochrane Central Register of Controlled Trials; National Library of Medicine's Medline database (PubMed); • Elsevier's Embase. Language restrictions: English, Italian, French and Spanish. Selection criteria Type of studies: systematic reviews, RCTs, CCTs, cross-sectional diagnostic studies, prospective or retrospective cohort studies, case series of at least 10 patients Participants: patients with breast cancer Intervention: FDG-PET or CT/PET Reference standard: histology or clinical follow up Comparator: any other imaging technique Outcomes> sensitivity, specificity, LR, metabolic/tumor response, time to recurrence, local, local-regional and distant recurrence, disease free survival, disease survival, overall survival Assessment of methodological quality of studies The following criteria have been used for the quality assessment of different study designs. Systematic reviews: criteria drawn from the AMSTAR checklist Diagnostic cross sectional studies: criteria drawn from the QUADAS checklist Randomized controlled trials: criteria suggested by the Cochrane Handbook

	Case control studies and cohort studies: criteria drawn from the New Castle-Ottawa checklist Level of evidence for estimates of diagnostic accuracy were assigned according to GRADE categorization of the quality of evidence Each member of the panel voted the level of appropriateness for each clinical question. Two rounds of votes were requested for the judgment of appropriateness and results were analysed using the RAND/UCLA Appropriateness Method. The use of FDG-PET for a specific clinical indication was judged was judged as *appropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 7-9 score region as *inappropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 1-3 score region. Finally the use of FDG-PET was judged as *uncertain* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 4-6 score region or when no agreement was reached after the second round of voting. Clinical indications for which the panel does not reach an agreement on level of appropriateness after two rounds of voting also fall in the *uncertain* category.
Conclusions	N STAGING OF PRIMARY BREAST CANCER - INAPPROPRIATE The panel agreed in judging as inappropriate the use of FDG-PET as a triage test, in order to identify patients eligible for axillary lymph node dissection, bypassing sentinel lymph node biopsy (SNLB). Level of evidence for diagnostic accuracy of FDG-PET resulted very low and the harm of an unnecessary axillary dissection was considered more severe than the benefit of bypassing SNLB. M STAGING OF LOCALLY ADVANCED BREAST CANCER - UNCERTAIN The panel did not reach an agreement in judging the role of FDG-PET in staging patients with locally advanced breast cancer (T3-T4 and/or N/N3) as a triage test, i.e. to direct FDG-PET positive patients to further more specific diagnostic tests. Level of evidence for diagnostic accuracy of FDG-PET was low, due partly to the heterogeneity of estimates for specificity, and ratings of panelists fell within all three regions (inappropriate, uncertain and appropriate). The final rating is therefore uncertain due to disagreement.
Notes	Meta-analysis of diagnostic accuracy estimates was not performed

SR - Cooper 2011 – N staging

Disease	breast cancer
Index test	FDG-PET, FDG-PET/CT
Comparators	MRI
Reference standard	histopathology following axillary lymph node dissection or sentinel lymph node biopsy
Target	diagnostic accuracy for axillary N staging
Studies included	diagnostic accuracy studies with prospective or retrospective set-up of study
Years covered by the search	up to April 2009
Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE, EMBASE, CINHAL, Cochrane Library, Science Citation Index, BIOSIS preview)
Characteristics of included studies clearly reported in tables	Yes (patients newly diagnosed with early-stage breast cancer [stage I, II, IIIA])
Methodological quality of primary studies assessed; criteria reported	Yes (QUADAS tool)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	FDG-PET and/or FDG-PET/CT: 26 studies (19 FDG-PET, 7 FDG-PET/CT) MRI: 9 studies
Design of included studies	FDG-PET and/or FDG-PET/CT: 15 prospective studies, 11 retrospective/unclear studies MRI: 6 prospective studies, 3 retrospective/unclear studies
N. of included patients	FDG-PET and/or FDG-PET/CT: 2591 patients (1729 FDG-PET, 862 FDG-PET/CT) MRI: 307 patients
Diagnostic accuracy results (with heterogeneity)	FDG-PET Sensitivity (pooled) 66.0% (95% CI 50.0-79.0%) Specificity (pooled) 93.0% (95% CI 89.0-96.0%) FDG-PET/CT

	Sensitivity (pooled) 56.0% (95% CI 44.0-67.0%) Specificity (pooled) 96.0% (95% CI 90.0-99.0%) MRI Sensitivity (pooled) 90% (95% CI 78.0-96.0%) Specificity (pooled) 90% (95% CI 75.0-96.0%)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Unclear (about 40% of studies with retrospective design) Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index

		test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

SR - Peare 2010 – N staging

Disease	breast cancer
Index test	FDG-PET, FDG-PET/CT
Comparators	clinical examination, ultrasound, MRI, mammography, MIBI
Reference standard	histopathology following axillary lymph node dissection or sentinel lymph node biopsy
Target	diagnostic accuracy for axillary N staging
Studies included	diagnostic accuracy studies with prospective or retrospective set-up of study
Years covered by the search	up to June 2009
Comprehensive bibliographic search: at least two databases searched	No (MEDLINE)

Characteristics of included studies clearly reported in tables	Yes
Methodological quality of primary studies assessed; criteria reported	Yes (prospective or retrospective set-up of study, recruitment consecutive or not, type of reference standard, independent reading of tests)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	No (not considered heterogeneity of estimates among studies)
N. of included studies	FDG-PET and/or FDG-PET/CT: 25 studies clinical examination: 7 studies ultrasound: 4 studies mammography: 2 studies MRI: 1 study MIBI: 1 study
Design of included studies	FDG-PET and/or FDG-PET/CT: 20 prospective studies, 5 retrospective/unclear studies
N. of included patients	FDG-PET and/or FDG-PET/CT: 2460 patients comparators: not reported number of patients
Diagnostic accuracy results (with heterogeneity)	FDG-PET or FDG-PET/CT Sensitivity (range) 20.0-100% Specificity (range) 66.0-100% Area under curve of the SROC curve 0.95 (CI 95% 0.91-0.97) Clinical examination Sensitivity (range) 40.0-60.0% Specificity (range) 85.0-100% Ultrasound Sensitivity (range) 52.0-100% Specificity (range) 83.0-100% MRI Sensitivity (1 study) 91.0% Specificity (1 study) 100%

	Mammography Sensitivity (range) 33.0-48.0% Specificity (range) 96.0-100% MIBI Sensitivity (1 study) 38.0% Specificity (1 study) 100%

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Yes (about 24% of studies with retrospective design) Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear (about 28% of studies with unclear or no blinding) Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

SR - Warning 2011 – N staging

Disease	breast cancer
Index test	FDG-PET, FDG-PET/CT
Comparators	none
Reference standard	not reported
Target	diagnostic accuracy for axillary N staging diagnostic accuracy for M staging (distant metastasis)
Studies included	diagnostic accuracy studies with prospective

	or retrospective set-up of study
Years covered by the search	up to 2010
Comprehensive bibliographic search: at least two databases searched	No (MEDLINE)
Characteristics of included studies clearly reported in tables	No
Methodological quality of primary studies assessed; criteria reported	No
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	No (not considered heterogeneity of estimates among studies)
N. of included studies	N staging FDG-PET: 25 studies FDG-PET/CT: 9 studies M staging FDG-PET or FDG-PET/CT: 6 studies
Design of included studies	not reported
N. of included patients	N staging FDG-PET: 2236 FDG-PET/CT: 859 M staging FDG-PET or FDG-PET/CT: 296
Diagnostic accuracy results (with heterogeneity)	N staging FDG-PET Sensitivity (range) 20.0-100% Specificity (range) 66.0-100% FDG-PET/CT Sensitivity (range) 20.0-98.0% Specificity (range) 84.0-100% M staging FDG-PET or FDG-PET/CT

	Sensitivity (range) 80.0-100% Specificity (range) 75.0-100%

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? Unclear Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear

		Did all patients receive a reference standard? Unclear Did patients receive the same reference standard? Unclear Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Unclear risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Unclear risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Unclear risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - N staging results

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Heudel 2010	FDG-PET/CT	45	patients with clinically resectable breast cancer	52.0 (95% CI 31.0–73.0%)	100 (95% CI 85.0–100%)
Pritchard 2012	FDG-PET/CT	325	patients with clinically resectable breast cancer (stage I,II)	23.7 (95% CI 15.9–33.6%)	99.6 (95% CI 97.2-99.9%)

SCC = squamous cell carcinoma

PS - Heudel 2010

Clinical features and settings	breast cancer; Country: France
Participants	Women with newly histologically proven breast cancer referred by surgeons for preoperative staging 45 patients; median age 55 years (range 26–85). Median tumor size 25 (range 8–90) mm; invasive ductal cancer in 37 of the 45 patients (82%) and invasive lobular cancer in 8 (18%). The prevalence of axillary node involvement at pathological examination was 51% (23 women)
Study design	diagnostic cross sectional study with prospective recruitment; after FDG-PET/CT imaging, all patients underwent breast surgery (mastectomy or breast-conserving surgery)
Target condition and reference standard(s)	axillary lymph nodes staging Reference standard: not reported (probably axillary lymph node dissection)
Index and comparator tests	FDG-PET/CT
Follow-up	not applicable
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Unclear Did patients receive the same reference standard? Unclear Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review

		question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Unclear risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Pritchard 2012

Clinical features and settings	breast cancer; Country: Canada
Participants	Women with resectable breast cancer and no clinical evidence of regional nodal or distant metastatic disease (stage I or II) 325 patients (out of 336 eligible); median age 56 years (range 28–83). Median tumor size 20 (range 1–90) mm. The prevalence of axillary node involvement at pathological examination was 29% (90 women). The prevalence of distant metastasis was 0.9% (3 women)
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	axillary lymph nodes staging Reference standard: axillary nodal assessment (histologic examination of resected axillary lymph nodes obtained by sentinel lymph node biopsy alone, by sentinel lymph node biopsy and axillary lymph node dissection, or by axillary lymph node dissection M staging (distant metastases) Reference standard: biopsy of positive lesions or clinical follow up
Index and comparator tests	FDG-PET/CT
Follow-up	not reported
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT	Low risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Yes

SELECTION (risk of bias)		Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? yes Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Yes Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? No, but low attrition (11 participants) Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting):

applicability)		Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

SR - Brennan 2012 – M staging

Disease	breast cancer
Index test	FDG-PET, FDG-PET/CT
Comparators	chest X-ray, abdominal ultrasound, bone scintigraphy, bone scan, chest and/or abdomen CT
Reference standard	clinical or imaging follow up, biopsy of positive lesions
Target	prevalence of asymptomatic metastatic disease diagnostic accuracy for M staging (distant metastases: bone, lung, liver metastases)
Studies included	diagnostic accuracy studies with prospective or retrospective set-up of study
Years covered by the search	up to July 2011
Comprehensive bibliographic search: at least two databases searched	No (only MEDLINE)
Characteristics of included studies clearly reported in tables	Yes (mixed population of stages and presentations at staging)
Methodological quality of primary studies assessed; criteria reported	Yes (prospective or retrospective set-up of study, recruitment consecutive or not, type od reference standard)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes: meta-analysis not performed due to the heterogeneity of clinical parameters across all studies
N. of included studies	22 studies only conventional imaging (abdominal ultrasound, chest X-ray, bone scan, CT, bone scintigraphy): 9 studies only FDG-PET and/or FDG-PET/CT: 8 studies both conventional imaging and FDG-PET and/or FDG-PET/CT: 5 studies all FDG-PET/CT: 6 studies
Design of included studies	only conventional imaging studies: 1 prospective study, 8 retrospective studies only FDG-PET and/or FDG-PET/CT: 5 prospective studies, 3 retrospective studies both conventional imaging and FDG-PET and/or FDG-PET/CT studies: 2 prospective studies, 3 retrospective

	studies
N. of included patients	only conventional imaging studies: 13860 patients only FDG-PET and/or FDG-PET/CT: 476 patients both conventional imaging and FDG-PET and/or FDG-PET/CT: 488 patients all FDG-PET/CT: 495 patients
Diagnostic accuracy results (with heterogeneity)	combined conventional imaging (abdomen ultrasound, chest X-ray, bone scintigraphy; 7 studies, 1299 participants) Sensitivity (median): 78.0% (range 33.3-100%) Specificity (median): 91.4% (range 67.3-97.9%) Bone scan (bone metastases) Sensitivity (median): 98% (range 33.3-100%) Specificity (median): 93.5% (range 85.4-100%) Chest X-ray (lung metastases) Sensitivity (median): 100% (range 40.0-100%) Specificity (median): 97.9% (range 96.4-99.0%) Liver ultrasound (liver metastases) Sensitivity (median): 100% (range 50.0-100%) Specificity (median): 96.7% (range 91.4-100%) chest and/or abdomen CT (lung and liver metastases; 5 studies 1470 participants) Sensitivity (median): 100% (range 87.0-100%) Specificity (median): 93.1% (range 85.7-97.6%) FDG-PET Sensitivity (median): 100% (range 78.0-100%) Specificity (median): 96.5% (range 82.0-100%) FDG-PET/CT Sensitivity (median): 100% (range 95.7-100%) Specificity (median): 98.1% (range 91.2-100%) Authors did not performed meta-analysis due to the heterogeneity of clinical parameters across studies

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	High risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? No (about 64% of studies with retrospective design) Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No

		Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	High risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

SR - Warning 2011 – M staging

Disease	breast cancer
Index test	FDG-PET, FDG-PET/CT
Comparators	none
Reference standard	not reported
Target	diagnostic accuracy for axillary N staging diagnostic accuracy for M staging (distant metastasis)
Studies included	diagnostic accuracy studies with prospective or retrospective set-up of study
Years covered by the search	up to 2010
Comprehensive bibliographic search: at least two databases searched	No (MEDLINE)
Characteristics of included studies clearly reported in tables	No
Methodological quality of primary studies assessed; criteria reported	No

Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	No (not considered heterogeneity of estimates among studies)
N. of included studies	N staging FDG-PET: 25 studies FDG-PET/CT: 9 studies M staging FDG-PET or FDG-PET/CT: 6 studies
Design of included studies	not reported
N. of included patients	N staging FDG-PET: 2236 FDG-PET/CT: 859 M staging FDG-PET or FDG-PET/CT: 296
Diagnostic accuracy results (with heterogeneity)	N staging FDG-PET Sensitivity (range) 20.0-100% Specificity (range) 66.0-100% FDG-PET/CT Sensitivity (range) 20.0-98.0% Specificity (range) 84.0-100% M staging FDG-PET or FDG-PET/CT Sensitivity (range) 80.0-100% Specificity (range) 75.0-100%

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? Unclear

		Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Unclear Did patients receive the same reference standard? Unclear Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Unclear risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review

		question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Unclear risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Unclear risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - M staging results

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Koolen 2012	FDG-PET/CT	154	patients eligible for neoadjuvant chemotherapy (stage II, III)	100	96.0
	conventional imaging			38.9	88.2
Pritchard 2012	FDG-PET/CT	325	patients with clinically resectable breast cancer (stage I,II)	100	96.8

PS - Koolen 2012

Clinical features and settings	breast cancer; Country: The Netherlands
Participants	Stage II and III breast cancer patients (= women who presented with invasive breast cancer >3 cm in diameter and/or at least one tumor-positive axillary lymph node) were offered to receive neoadjuvant chemotherapy in our institute. Prior to the start of chemotherapy, a search for distant disease was performed with whole body 18F-FDG PET/CT as well as with conventional imaging techniques (bone scintigraphy, ultrasound of the liver, and chest radiography) 167 eligible patients. Conventional staging was not complete in 13 patients, resulting in 154 patients included into analysis (mean age ± SD 49.1 ± 11.0) N-stage prior to neoadjuvant chemotherapy cN0 43 (28%) cN1 83 (54%) cN2 4 (3%) cN3 24 (16%)
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	distant metastasis at staging Reference standard: confirmation of suspect lesions obtained by cytological or histological verification or, if not available or possible, with additional imaging studies or by prolonged follow-up.
Index and comparator tests	FDG-PET/CT; comparators: conventional imaging techniques (bone scintigraphy, ultrasound of the liver, and chest radiography)
Follow-up	median 9.0 months (range 6.6–24.6 months)

Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	High risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? No If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? No Did patients receive the same reference standard? Unclear

		Were all patients included in the analysis? No Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Pritchard 2012

Clinical features and settings	breast cancer; Country: Canada
Participants	Women with resectable breast cancer and no clinical evidence of regional nodal or distant metastatic disease (stage I or II) 325 patients (out of 336 eligible); median age 56 years (range 28–83). Median tumor size 20 (range 1–90) mm. The prevalence of axillary node involvement at pathological examination was 29% (90 women). The prevalence of distant metastasis was 0.9% (3 women)
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	axillary lymph nodes staging Reference standard: axillary nodal assessment (histologic examination of resected axillary lymph nodes obtained by sentinel lymph node biopsy alone, by sentinel lymph node biopsy and axillary lymph node dissection, or by axillary lymph node dissection M staging (distant metastases) Reference standard: biopsy of positive lesions or clinical follow up
Index and comparator tests	FDG-PET/CT

Follow-up	not reported
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Yes Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes

		Did patients receive the same reference standard? No Were all patients included in the analysis? No, but low attrition (11 participants) Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

Characteristics of excluded studies

Ahn 2010

Reason for exclusion	index test: only FDG-PET

Berg 2011

Reason for exclusion	study on positron emission mammography

Berg 2012

Reason for exclusion	study on positron emission mammography

Bruening 2012

Reason for exclusion	target condition: primary tumor diagnosis

Carkaci 2012

Reason for exclusion	retrospective study

Choi 2011

Reason for exclusion	retrospective study

Chu 2012

Reason for exclusion	prognostic study

Escalona 2010

Reason for exclusion	systematic review included in former HTA report

Fosse 2012

Reason for exclusion	retrospective study

Garami 2012

Reason for exclusion	unclear if prospective study

Gilardi 2010

Reason for exclusion	population: restaging after neoadjuvant therapy

Grankvist 2012

Reason for exclusion	target condition: suspected recurrence

Hahn 2011

Reason for exclusion	retrospective study

Heusner 2010

Reason for exclusion	patients with suspected recurrence 40% of whole sample

Houssami 2011

Reason for exclusion	target condition: suspected recurrence or restaging

Kim 2012

Reason for exclusion	prognostic retrospective study

Kong 2010

Reason for exclusion	retrospective study

Liu 2011

Reason for exclusion	target condition: suspected recurrence or restaging

Mittal 2011

Reason for exclusion	retrospective study

Morris 2010

Reason for exclusion	retrospective study

Moy 2010

Reason for exclusion	target condition: primary tumor diagnosis

Niikura 2011

Reason for exclusion	retrospective study

Niikura 2011a

Reason for exclusion	retrospective study

Pan 2010

Reason for exclusion	systematic review included in former HTA report

Piccardo 2012

Reason for exclusion	target condition: suspected recurrence

Robertson 2011

Reason for exclusion	narrative review

Schilling 2011

Reason for exclusion	study on positron emission mammography

Segaert 2010

Reason for exclusion	retrospective study

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”:ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw or
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. “Breast neoplasm” MeSH descriptor
10. “Carcinoma, Ductal, Breast” MeSH descriptor
11. “Phyllodes Tumor” MeSH descriptor
12. breast NEAR (tumor* OR cancer* OR neoplasm*): ti,ab,kw
13. Mammary NEAR (neoplasm * or carcinoma*): ti,ab,kw
14. Philloides: ti,ab,kw
15. 10/14 OR
16. 8 AND 15

Publication date: January 2010 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18” [Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose” [All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg*[All Fields]
14. fdg 18*[All Fields]
15. fdg/*[All Fields]
16. “fdg pet” [All Fields]
17. “Positron-Emission Tomography” [Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/24 OR
25. “Breast Tumor” [title/abstract]
26. “Breast Cancer” [title/abstract]
27. “Mammary Carcinoma” [title/abstract]
28. “breast neoplasm” [title/abstract]
29. “breast neoplasms” [title/abstract]
30. “Mammary Neoplasm” [title/abstract]
31. “Breast Neoplasms” [Mesh: NoExp]
32. “Carcinoma, Ductal, Breast” [Mesh]
33. “Mammary Ductal Carcinoma” [ti/ab]
34. “Phyllodes Tumor” [Mesh: NoExp]
35. “Phyllodes” [titile/abstract]
36. 25/36 OR
37. 24 AND 36
38. “editorial” [Publication Type]
39. “comment” [Publication Type]

40. “letter” [Publication Type]
41. “review” [Publication Type]
42. “case reports” [Publication Type]
43. 38/42 OR
44. 37 NOT 43

Limits: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2010 - March 2012

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “fluorodeoxyglucose f 18”/exp
3. (“fluorodeoxyglucose f 18”/syn
4. “computer assisted emission tomography”/exp
5. “computer assisted emission tomography” OR
6. pet
7. “pet scans”
8. “pet scanner”
9. “pet scan”
10. “pet/ct scan”
11. “pet/ct scans”
12. “pet/ct”
13. “positron emission tomography/computed tomography”
14. pet NEAR/4 scan*
15. pet NEAR/4 ct
16. 1/15 OR
17. “breast cancer”/syn
18. “breast cancer”
19. “breast neoplasm”
20. mammary NEAR/2 carcinoma
21. “breast sarcoma”
22. “breast adenocarcinoma”
23. phyllodes
24. “inflammatory breast cancer”
25. “intraductal carcinoma”
26. “ductal carcinoma”
27. “paget breast disease”
28. “breast cancer”/de
29. “breast adenocarcinoma”/exp
30. “breast carcinoma”/exp
31. “breast metastasis”/exp
32. “breast sarcoma”/exp
33. cystosarcoma phylloides”/exp
34. inflammatory breast cancer”/exp
35. “intraductal carcinoma”/exp
36. “paget nipple disease”/exp
37. 17/36
38. 37 AND 16

Limits: Humans
Languages: English, French, Italian, Spanish
Publication date: July 2010 - March 2012

APPENDIX 7

FDG-PET/CT for staging of esophageal cancer

HTA report - ASSR 2011 esophageal cancer

Document ID	HTA report - ASSR 2011 esophageal cancer
Objectives	to define criteria for appropriate use of FDG-PET for patients with esophageal cancer
Methods	A panel of experts working in Health Trusts and Teaching Hospitals of Emilia-Romagna was convened to discuss and agree on the methodology for a research programme aimed at defining the criteria for appropriate use of PET in esophageal cancer. On the basis of the clinical pathway of patients with esophageal cancer the panel examined and assessed the role of FDG-PET for 7 clinical indications (N staging of primary esophageal cancer; M staging of primary esophageal cancer; target volume definition of curative radiation treatment; evaluation of early response to neoadjuvant therapy; evaluation of response to neoadjuvant therapy at the end of treatment; follow up in patients with no suspicion of recurrence; diagnosis and staging of suspect distant recurrence). The following databases were searched for the period between January 2006 and July 2010: Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE - The Cochrane Library); Health Technology Assessment Database; Cochrane Central Register of Controlled Trials; National Library of Medicine's Medline database (PubMed); Elsevier's Embase. Language restrictions: English, Italian, French and Spanish. Selection criteria Type of studies: systematic reviews, RCTs, CCTs, cross-sectional diagnostic studies, prospective or retrospective cohort studies, case series of at least 10 patients Participants: patients with esophageal cancer Intervention: FDG-PET or CT/PET Reference standard: histology or clinical follow up Comparator: any other imaging technique Outcomes> sensitivity, specificity, LR, metabolic/tumor response, time to recurrence, local, local-regional and distant recurrence, disease free survival, disease survival, overall survival Assessment of methodological quality of studies The following criteria have been used for the quality assessment of different study designs. Systematic reviews: criteria drawn from the AMSTAR checklist Diagnostic cross sectional studies: criteria drawn from the QUADAS checklist Randomized controlled trials: criteria suggested by the Cochrane Handbook Case control studies and cohort studies: criteria drawn from the New Castle-Ottawa checklist Level of evidence for estimates of diagnostic accuracy were assigned according to GRADE categorization of the quality of evidence Each member of the panel voted the level of appropriateness for each clinical question. Two

	rounds of votes were requested for the judgment of appropriateness and results were analysed using the RAND/UCLA Appropriateness Method. The use of FDG-PET for a specific clinical indication was judged was judged as *appropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 7-9 score region as *inappropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 1-3 score region. Finally the use of FDG-PET was judged as *uncertain* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 4-6 score region or when no agreement was reached after the second round of voting. Clinical indications for which the panel does not reach an agreement on level of appropriateness after two rounds of voting also fall in the *uncertain* category.;
Conclusions	N STAGING OF PRIMARY ESOPHAGEAL CANCER - UNCERTAIN The panel agreed to judge as uncertain the use of FDG-PET in staging patients with esophageal cancer for regional lymph nodes, in replacement of endoscopic ultrasonography (EUS). The level of evidence for diagnostic accuracy of FDG-PET was very low, with heterogeneous estimates for both sensitivity and specificity. All outcomes, related to the correct selection of patients eligible for neoadjuvant chemoradiation therapy were considered “important” (median score 6). A less invasive test was also deemed highly desirable, given the high pre-test probability of patients diagnosed for primary esophageal cancer having positive lymph node. However the uncertainty on the diagnostic accuracy of FDG-PET made the panel very cautious in suggesting the use of FDG-PET results to direct therapeutic options. M STAGING OF PATIENTS AND DETECTION OF SYNCHRONOUS SECOND PRIMARY TUMOR IN PATIENTS WITH LOCALLY ADVANCED HEAD AND NECK CANCER - APPROPRIATE At the first voting round the panel agreed to judge appropriate the use of FDG-PET for M staging of advanced head and neck cancer in patients with negative or equivocal results from conventional imaging. Level of evidence for diagnostic accuracy of FDG-PET was judged moderate with estimates for sensitivity higher than conventional imaging. All clinical outcomes were considered “critical” (median score 8), with a closer range (between 7 and 8) for patients correctly upstaged, highlighting the added value of FDGPET in identifying patients with distant metastases or second primary tumors missed by conventional imaging.
Notes	Meta-analysis of diagnostic accuracy estimates was not performed

PS - N staging results

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Hsu 2011	FDG-PET/CT	76	squamous cell carcinoma	52.4	87.3

PS - Hsu 2011

Clinical features and settings	esophageal cancer; Country: Taiwan
Participants	Patients without distant metastasis or definite evidence of extensive adjacent organ invasion undergoing surgical resection. 125 patients eligible; only 76 (63 males and 13 females) included (those treated with the triincisional approach, which included right thoracotomy, midline laparotomy, and left cervicotomy, or video-assisted thoracoscopic esophagectomy; those having squamous cell carcinoma; those assessed with FDG-PET/CT); mean age 61.7 years (SD 10.9); squamous cell carcinomas 100% Tumor invasion depth T1: 18 (23.7 %) T2: 7 (9.2%) T3: 49 (64.5%) T4: 2 (2.6%)
Study design	diagnostic cross sectional study with prospective recruitment; patients undergoing esophagectomy after diagnostic work-up
Target condition and reference standard(s)	regional lymph nodes staging (N2 or N3 status versus N1 or N0 status); Reference standard: postoperative pathologic staging
Index and comparator tests	FDG-PET/CT; comparators: none
Follow-up	not applicable
Notes	Results regional lymph nodes staging FDG-PET/CT sensitivity: 52.4% specificity: 87.3%

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	High risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard?

		Yes Were all patients included in the analysis? No Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? Only squamous cell carcinomas CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

Characteristics of excluded studies

Abdelsalam 2010

Reason for exclusion	not FDG-PET/CT

Aigner 2010

Reason for exclusion	abstract at congress

Aigner 2011

Reason for exclusion	full-text not available

Alam 2011

Reason for exclusion	abstract at congress

Alan 2010

Reason for exclusion	abstract at congress

Aoyagi 2010

Reason for exclusion	not a pertinent research question

Attia 2011

Reason for exclusion	abstract at congress

Barber 2011

Reason for exclusion	abstract at congress

Blom 2011

Reason for exclusion	not a pertinent research question

Blom 2011a

Reason for exclusion	sensitivity and specificity estimates not available

Chan 2011

Reason for exclusion	not esophageal cancer

Chen 2011

Reason for exclusion	not question on staging

Choi 2010

Reason for exclusion	abstract at congress

Choi 2010a

Reason for exclusion	included in ASSR 2011 HTA report

Crabtree 2011

Reason for exclusion	retrospective study

De Vita 2010

Reason for exclusion	abstract at congress

Eloubeidi 2011

Reason for exclusion	retrospective study

Gillies 2011

Reason for exclusion	not a diagnostic accuracy question

Goenka 2011

Reason for exclusion	abstract at congress

Guo 2010

Reason for exclusion	abstract at congress

Han 2011

Reason for exclusion	per-node analysis

Jung 2010

Reason for exclusion	abstract at congress

Kaida 2010

Reason for exclusion	abstract at congress

Karashima 2010

Reason for exclusion	abstract at congress

Kayani 2011

Reason for exclusion	not question on staging

Marzola 2012

Reason for exclusion	narrative review

Monjazeb 2010

Reason for exclusion	response to treatment question

Natsugoe 2010

Reason for exclusion	abstract at congress

Okazumi 2010

Reason for exclusion	abstract at congress

Peng 2010

Reason for exclusion	abstract at congress

Schreurs 2011

Reason for exclusion	not FDG-PET/CT

Shan 2010

Reason for exclusion	abstract at congress

Shuto 2010

Reason for exclusion	abstract at congress

Sohda 2010

Reason for exclusion	retrospective study

Sosef 2010

Reason for exclusion	abstract at congress

Staiger 2010

Reason for exclusion	full text not available

Sun 2011

Reason for exclusion	abstract at congress

Syed 2011

Reason for exclusion	abstract at congress

Tanabe 2011

Reason for exclusion	retrospective study

Thurau 2011

Reason for exclusion	not question on staging

van Heijl 2010

Reason for exclusion	included in ASSR 2011 HTA report

van Heijl 2011

Reason for exclusion	not question on staging

Walker 2011

Reason for exclusion	sensitivity and specificity estimates not available

Wilson 2010

Reason for exclusion	abstract at congress

Won 2010

Reason for exclusion	abstract at congress

Wong 2012

Reason for exclusion	guidelines document

Yasuda 2012

Reason for exclusion	retrospective study

Yen 2012

Reason for exclusion	retrospective study

Yu 2011

Reason for exclusion	per-node analysis

Zhong 2010

Reason for exclusion	abstract at congress

Zhu 2010

Reason for exclusion	abstract at congress

zum Buschenfelde 2011

Reason for exclusion	not question on staging

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”: ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw or
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. “Esophageal Neoplasms” MeSH descriptor explode all trees
10. “esophageal cancer”: ti,ab,kw
11. “Esophageal Neoplasm”: ti,ab,kw
12. “Esophagus Cancer”: ti,ab,kw
13. “Esophagus Neoplasm”: ti,ab,kw
14. “esophageal cancer”: ti,ab,kw
15. “esophagus cancer”: ti,ab,kw
16. 10/15 OR
17. 8 AND 16

Publication date: July 2010 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18”[Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18fdg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg pet”[All Fields]
17. “Positron-Emission Tomography”[Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/24 OR
25. “esophageal cancer”[All Fields]
26. “esophagus cancer”[All Fields]
27. “Esophageal Neoplasms”[Mesh]
28. “Esophageal Neoplasm”
29. “Esophageal Cancer”
30. “Esophagus Neoplasm”
31. “Esophagus Cancer”
32. 25/31 OR
33. 24 AND 32
34. “editorial”[Publication Type]
35. “comment”[Publication Type]
36. “letter”[Publication Type]
37. “review”[Publication Type]
38. “case reports”[Publication Type]
39. 34/38 OR

40. 33 NOT 39
Limits: Humans

Publication date: July 2010 - March 2012

3 EMBASE search strategy

1. ‘esophagus cancer’/exp
2. ‘esophagus cancer’
3. ‘esophagus cancer’/syn
4. ‘esophageal NEXT (cancer OR neoplasm OR tumor)
5. 1/4 OR
6. ‘positron emission tomography’/syn
7. ‘fluorodeoxyglucose f 18’/exp
8. (‘fluorodeoxyglucose f 18’/syn
9. ‘computer assisted emission tomography’/exp
10. ‘computer assisted emission tomography’ OR
11. pet
12. ‘pet scans’
13. ‘pet scanner’
14. ‘pet scan’
15. ‘pet/ct scan’
16. ‘pet/ct scans’
17. ‘pet/ct’
18. ‘positron emission tomography/computed tomography’
19. pet NEAR/4 scan*
20. pet NEAR/4 ct
21. 6/20 OR
22. 5 AND 21

Limits: Humans

Publication date: July 2010 - March 2012

APPENDIX 8

FDG-PET/CT for staging of stomach cancer

Characteristics of included studies

HTA report - KCE 2009

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients; • Absence of adequate reference standard;

	• Absence of patient-based analysis; • Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	no systematic reviews or primary studies were found regarding gastric cancer staging.
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instruments which are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

SR - Kwee 2009 - N staging

Disease	**gastric cancer (adenocarcinoma)**
Index test	FDG-PET (4 studies) or FDG-PET/CT (one study)
Comparators	abdominal ultrasonography (AUS) endoscopic ultrasonography (EUS) multidetector-row computed tomography (MDCT) magnetic resonance imaging (MRI)
Reference standard	histopathological examination after surgery or clinical follow-up
Target	detection of lymph node metastases (N-staging)
Studies included	FDG-PET as index test: 4 studies included (Mukai 2006, Yun 2005, Tian 2004, Yeung 1998); FDG-PET/CT as index test: 1 study included (Yang 2008).
Years covered by the search	No beginning date limit was used. The search was updated until July 7, 2008.
Comprehensive bibliographic search: at least two databases searched	YES: PubMed/MEDLINE and Embase databases
Characteristics of included studies clearly reported in tables	YES
Methodological quality of primary studies assessed; criteria reported	YES: The methodological quality of the included studies was assessed in terms of the potential for bias (internal validity) and lack of generalizability (external validity) according to a modified-QUADAS tool. For each of the included studies, 13 methodological quality items were assessed (maximum total score: 100%; a study was judged of high quality if score > 60%). For the FDG-PET studies, the total methodological quality score ranged from 46% to 62% (median, 58%). Two FDG-PET studies (Mukai 2006 and Yun 2005) were of high methodological quality. For the only FDG-PET/CT study (Yang 2008), the total methodological quality score was 54%.
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	No, because of the hign level of heterogeneity and moderate methodological quality of the included studies.

N. of included studies	4 studies included for FDG-PET and 1 study for FDG-PET/CT abdominal ultrasonography (AUS) 6 studies endoscopic ultrasonography (EUS) 30 studies multidetector-row computed tomography (MDCT) 10 studies magnetic resonance imaging (MRI) 3 studies
Design of included studies	FDG-PET: among the 4 included studies, only one (Tian 2004) was prospective. FDG-PET/CT: the only included study was retrospective.
N. of included patients	FDG-PET: 183 patients (Mukai 2006: 62; Yun 2005: 81; Tian 2004: 27; Yeung 1998: 13). FDG-PET/CT: 78 patients (Yang 2008).
Diagnostic accuracy results (with heterogeneity)	The sensitivity and specificity of AUS for the detection of lymph node metastasis varied between 12.2% and 80.0% (median, 39.9%) and 56.3% and 100% (median, 81.8%). The sensitivity and specificity of EUS varied between 16.7% and 96.8% (median, 70.8%) and 48.4% and 100% (median, 84.6%). The sensitivity and specificity of MDCT varied between 62.5% and 91.9% (median, 80.0%) and 50.0% and 87.9% (median, 77.8%). The sensitivity and specificity of MRI varied between 54.6% and 85.3% (median, 68.8%) and 50.0% and 100% (median, 75.0%). The sensitivity and specificity of FDG-PET varied between 33.3% and 64.6% (median, 34.3%) and 85.7% and 97.0% (median, 93.2%) respectively. There was no significant difference between the mean sensitivity of FDG-PET studies with high and low methodological quality (34.3% vs 49.0%; P = 0.515). There also was no significant difference between the mean specificity of studies with high and low methodological quality (96.7% vs 87.9%; P = 0.131). The sensitivity and specificity of the only one FDG-PET/CT study included were 54.7% and 92.2%, respectively.
Notes: The aim of this study was to systematically review the current role of imaging (FDG-PET, FDG-PET/CT, abdominal ultrasonography (AUS), endoscopic ultrasonography (EUS), multidetector-row computed tomography (MDCT), magnetic resonance imaging (MRI) in assessing lymph node (LN) status in gastric cancer.	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	High risk	Describe methods of patient selection: ϖ Was a consecutive or random sample of patients enrolled? **No** (Yes only for Mukai 2006) ϖ Was a case-control design avoided? **Unclear** ϖ Did the study avoid inappropriate exclusions? **Unclear** Could the selection of patients have introduced bias? CONCERN: LOW=YES/**HIGH=NO**/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: ϖ Were the index test results interpreted without knowledge of the results of the reference standard? **Unclear** (blinding there was in 2/4 studies FDG-PET and in the study FDG-PET/CT) ϖ If a threshold was used, was it pre-specified? **Unclear** Could the conduct or interpretation of the index test have introduced bias? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**
3A. REFERENCE STANDARD (risk of bias)	Low risk	Describe the reference standard and how it was conducted and interpreted: ϖ Is the reference standard likely to correctly classify the target condition? **Yes** ϖ Were the reference standard results interpreted without knowledge of the results of the index test? **Unclear** Could the reference standard, its conduct, or its interpretation have introduced bias? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: ϖ Was there an appropriate interval between index test(s) and reference standard? **Unclear** ϖ Did all patients receive a reference standard? **Unclear** ϖ Did patients receive the same reference standard? **Unclear** ϖ Were all patients included in the analysis? **Unclear** Could the patient flow have introduced bias?

		CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? **LOW=YES**/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? **LOW=YES**/HIGH=NO/UNCLEAR

SR - Wang 2011 - FDG-PET for M staging

Disease	gastric cancer (adenocarcinoma)
Index test	FDG-PET
Comparators	ultrasonography(US), endoscopic ultrasound (EUS), computed tomography (CT), magnetic resonance imaging (MRI)
Reference standard	histopathological examination after surgery or clinical follow-up
Target	detection of hepatic and peritoneal metastases (M-staging)
Studies included	diagnostic accuracy primary studies (study design element as "prospective or retrospective" was not reported).
Years covered by the search	No beginning date limit was used. The search was updated until February , 2011.
Comprehensive bibliographic search: at least two databases searched	YES: Pubmed/Medline, Embase, The Cochrane Library and the China Biological Medicine Databases.
Characteristics of included studies clearly reported in tables	YES
Methodological quality of primary studies assessed; criteria reported	YES: The methodological quality of the included studies was assessed according to QUADAS tool.
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	US 8 studies; EUS 5 studies; CT 22 studies; MRI 2 studies 5 studies included for FDG-PET (Yoshioka 2003; Yeung 1998; Yun 2005; Lim 2006;Chen 2005)
Design of included studies	diagnostic accuracy primary studies but study design element as "prospective or retrospective" was not reported.
N. of included patients	FDG-PET: 338 patients

Diagnostic accuracy results (with heterogeneity)	**Liver metastases** US Pooled sensitivity 0.54 (95% CI: 0.34-0.73) Pooled specificity 0.98 (95% CI: 0.90-0.99) CT Pooled sensitivity 0.74 (95% CI: 0.59-0.85). Pooled specificity 0.99 (95% CI: 0.97-1.00) Only two studies' data were sufficient for EUS and MRI, so pooled analysis was not conducted FDG-PET: data were available for four studies. Pooled sensitivity and specificity in detecting liver metastasis was 0.70 (95% CI: 0.36-0.90) and 0.96 (95% CI: 0.81-0.99), respectively. **Peritoneal metastases** US Pooled sensitivity 0.09 (95% CI: 0.03-0.21) Pooled specificity 0.99 (95% CI: 0.96-1.00) EUS Pooled sensitivity 0.34 (95% CI: 0.10-0.69) Pooled specificity 0.96 (95% CI: 0.87-0.99) CT Pooled sensitivity 0.33 (95% CI: 0.16-0.56) Pooled specificity 0.99 (95% CI: 0.98-1.00) FDG-PET: data were available for four studies. Pooled sensitivity and specificity was 0.28 (95% CI: 0.17-0.44) and 0.97 (95% CI: 0.83-1.00), respectively. Meta-analysis was based on the bivariate model in the presence of significant heterogeneity.
Notes: The aim of this study was to systematically review the current role of imaging (FDG-PET, ultrasonography (US), endoscopic ultrasonography (EUS), computed tomography (CT), magnetic resonance imaging (MRI) in assessing hepatic and peritoneal metastases in gastric cancer; each imaging technology has been reviewed as index test, separately.	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: ϖ Was a consecutive or random sample of patients enrolled? **Yes** ϖ Was a case-control design avoided? **Unclear** ϖ Did the study avoid inappropriate exclusions? **No** Could the selection of patients have introduced bias? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: ϖ Were the index test results interpreted without knowledge of the results of the reference standard? **Unclear** ϖ If a threshold was used, was it pre-specified? **Unclear** Could the conduct or interpretation of the index test have introduced bias? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: ϖ Is the reference standard likely to correctly classify the target condition? **Unclear** ϖ Were the reference standard results interpreted without knowledge of the results of the index test? **Unclear** Could the reference standard, its conduct, or its interpretation have introduced bias? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**
4. FLOW AND TIMING (risk of bias)	High risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: ϖ Was there an appropriate interval between index test(s) and reference standard? **Unclear** ϖ Did all patients receive a reference standard? **No** ϖ Did patients receive the same reference standard? **No** ϖ Were all patients included in the analysis? **Unclear**

		Could the patient flow have introduced bias? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=**YES**/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=**YES**/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=**YES**/HIGH=NO/UNCLEAR

PS - Chung 2010 – M staging

Clinical features and settings	gastric cancer (adenocarcinoma); country: Korea
Participants	35 consecutive newly diagnosed patients with gastric carcinoma underwent FDG-PET/CT during the period from April 2006 to December 2008 were included; mean age (range): 57 ± 13 (31-82); male 24.
Study design	diagnostic accuracy measurement (within a prognostic cohort study)
Target condition and reference standard(s)	Solid organ metastases (lung, liver, bone, or adrenal gland) were assessed as distinct from peritoneum or nonregional lymph node metastases (M Staging); Reference standard: histologic confirmation or by contrast-enhanced CT and serial follow-up.
Index and comparator tests	FDG-PET/CT; comparators: CT, bone scintigraphy, magnetic resonance imaging.
Follow-up	Follow up monitoring for recurrence or metastasis was performed every 2-3 months
Notes	Diagnostic test accuracy measures are extracted from a prognostic cohort study

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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? **No** Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? **Unclear** If a threshold was used, was it pre-specified? **Yes** Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target

		condition? **Unclear** Were the reference standard results interpreted without knowledge of the results of the index test? **Unclear** Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? **Unclear** Did all patients receive a reference standard? **Unclear** Did patients receive the same reference standard? **No** Were all patients included in the analysis? **Yes** Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
1.B PATIENT SELECTION (concern of applicability)	Unclear risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Unclear risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES**/**HIGH=NO**/UNCLEAR**

Characteristics of excluded studies

Cordin 2010

Reason for exclusion	narrative review

Hiraoka 2010

Reason for exclusion	FDG-PET/CT is the reference standard.

Hur 2010

Reason for exclusion	retrospective study

Kim 2011

Reason for exclusion	retrospective study

Roedl 2009

Reason for exclusion	case-control study

Saif 2010

Reason for exclusion	pet-ct overview in oncology

Shimada 2011

Reason for exclusion	narrative review

Smyth 2011

Reason for exclusion	overview

Sun 2010

Reason for exclusion	retrospective study

Suttie 2009

Reason for exclusion	predictive studies review

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”: ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. Stomach Neoplasms [Mesh explodes all trees]
10. Stomach Neoplasm*: ti,ab,kw
11. Stomach Cancer* : ti,ab,kw
12. Gastric Cancer*: ti,ab,kw
13. Gastric Neoplasm*:ti,ab,kw
14. Gastric NEAR/4 cancer*: ti,ab,kw
15. Gastric NEAR/4 neoplasm*: ti,ab,kw
16. Stomach NEAR/4 cancer*: ti,ab,kw
17. Stomach NEAR/4 cancer*: ti,ab,kw
18. 9/17 OR
19. 8 AND 18

Publication date: January 2009 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18” [Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg-pet”[All Fields]
17. “Positron-Emission Tomography” [Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. Stomach Neoplasms [Mesh explodes all trees]
26. “Stomach Neoplasm” [Title/Abstract]
27. “Stomach Neoplasms” [Title/Abstract]
28. “Gastric Neoplasm” [Title/Abstract]
29. “Gastric Neoplasms” [Title/Abstract]
30. “Stomach Cancer” [Title/Abstract]
31. “Stomach Cancers” [Title/Abstract]
32. “Gastric Cancer” [Title/Abstract]
33. “Gastric Cancers” [Title/Abstract]
34. 25/3 OR

35. 24 AND 34

Limit: Humans

Publication date: January 2009 - March 2012

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “positron emission tomography”/exp
3. “fluorodeoxyglucose f 18”/exp
4. “fluorodeoxyglucose f 18”/syn
5. “computer assisted emission tomography”/exp
6. “computer assisted emission tomography”/tw
7. pet/tw
8. “pet scans”/tw
9. “pet scanner”/tw
10. “pet scan”/tw
11. “pet/ct scan”/tw
12. “pet/ct scans”/tw
13. “pet/ct”/tw
14. “positron emission tomography/computed tomography”/tw
15. pet NEAR/4 scan*
16. pet NEAR/4 ct
17. 1/15 OR
18. “Stomach Neoplasms”/de, syn, Keyword
19. “Stomach Neoplasms”/exp
20. “stomach cancer”/de, syn, Keyword”
21. “stomach cancers”/de, syn, Keyword”
22. “gastric Neoplasms”/de, syn, Keyword
23. “gastric cancer”/de, syn, Keyword
24. “gastric cancers”/de,syn;keywordOR“gastric neoplasm”/de,syn, keyword
25. “stomach cancer”: ti, ab.
26. “stomach neoplasm”: ab:ti
27. “stomach neoplasms” : ab:ti
28. “stomach cancers”: :ab:ti
29. “gastric cancers”: :ab:ti
30. “gastric cancer”:ab:ti
31. “gastric neoplasm”:ab:ti
32. “gastric neoplasms”:ab:ti
33. Stomach NEAR/4 neoplasm*
34. Stomach NEAR/4 cancer*
35. Gastric NEAR/4 neoplasm*
36. Gastric NEAR/4 cancer*
37. 18/36 OR
38. 17 AND 37

Limit: Humans

Limit: “article” OR “review”/it OR “short survey”

Publication date: January 2009 - March 2012

APPENDIX 9

FDG-PET/CT for staging of pancreatic cancer

Characteristics of included studies

HTA report - KCE 2009 pancreatic cancer

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients; • Absence of adequate reference standard; • Absence of patient-based analysis; • Case-control study design;

	• Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	The 2009 KCE report assessed the use of FDG-PET/CT and FDG-PET for diagnostic and staging purposes. On the basis of 13 retrospective and prospective studies, KCE reported finding limited evidence of diagnostic utility and similarly limited evidence of sensitivity and specificity of the test for staging. For both indications it found no evidence of benefit compared to the currently available alternatives CT and EUS, US and ERCP. The conclusions were that the utility of FDG PET/CT for both diagnosis and staging needs confirmation from further larger studies.
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

Characteristics of excluded studies

Abgral 2011

Reason for exclusion	Study comparing PET with scintigraphy

Buchs, 2011

Reason for exclusion	Diagnosis of primary cancer as target condition

Herrmann 2012

Reason for exclusion	Fluorothymidine was used as contrast medium

Kauhanen 2009

Reason for exclusion	Diagnosis of primary cancer as target condition

Kitajima 2010

Reason for exclusion	Study assessing diagnostic accuracy of PET/CT in re-staging

Kuwatani 2009

Reason for exclusion	Study comparing with CT and markers

Okano 2011

Reason for exclusion	Participants were recruited retrospectively

Tang 2011

Reason for exclusion	Searches to April 2009 but no new studies

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

“Positron-Emission Tomography” [MeSH descriptor explode all trees] OR “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees] OR. “positron emission tomography”: ti,ab,kw OR. pet*: ti,ab,kw OR. pet scan*: ti,ab,kw OR. “Fluorodeoxyglucose F18”: ti,ab,kw OR. fdg NEAR/2 18: ti,ab,kw	AND	Pancreatic Neoplasms [Mesh **explodes all trees] contiene** Neoplasm, Pancreatic **Pancreatic Neoplasm** Neoplasms, Pancreatic Pancreas Neoplasms Neoplasm, Pancreas Neoplasms, Pancreas Pancreas Neoplasm Cancer of Pancreas Pancreas Cancers Pancreas Cancer Cancer, Pancreas Cancers, Pancreas Pancreatic Cancer Cancer, Pancreatic Cancers, Pancreatic Pancreatic Cancers Cancer of the Pancreas OR Pancreatic Neoplasm*: ti,ab,kw OR Pancreatic Cancer* : ti,ab,kw OR “Cancer* of Pancreas” : ti,ab,kw OR Pancreatic NEAR/4 cancer*: ti,ab,kw OR Pancreatic NEAR/4 neoplasm*: ti,ab,kw OR “Pancreatic Adenocarcinoma”: ti,ab,kw

2 MEDLINE search strategy

"Fluorodeoxyglucose F18"[Mesh] **OR** "2-Fluoro-2-deoxyglucose" [All Fields] **OR** "18F Fluorodeoxyglucose" [All Fields] **OR** "F 18 Fluorodeoxyglucose" [All Fields] **OR** Fludeoxyglucose* [All Fields] **OR** "2 fluoro 2 deoxy d glucose"[All Fields] **OR** 18fluorodesoxyglucose*[All Fields] **OR** fluorodeoxyglucose*[All Fields] **OR** "fluorine 18 fluorodeoxyglucose" [All Fields] **OR** 18f dg*[All Fields]) **OR** 18fluorodeoxyglucose*[All Fields] **OR** 18fdg [All Fields] **OR** 18 fdg* [All Fields] **OR** fdg 18* [All Fields] **OR** fdg/* [All Fields] **OR** "fdg pet"[All Fields] **OR** "Positron-Emission Tomography"[Mesh] **OR** "positron emission tomography" [title/abstract] **OR** pet [title/abstract] **OR** "pet scan" [All Fields] **OR** "pet scans" [All Fields] **OR** "pet scanner" [All Fields] **OR** petscan [All Fields]	AND	Pancreatic Neoplasms [Mesh **explodes all trees] contiene** Neoplasm, Pancreatic **Pancreatic Neoplasm** Neoplasms, Pancreatic Pancreas Neoplasms Neoplasm, Pancreas Neoplasms, Pancreas Pancreas Neoplasm Cancer of Pancreas Pancreas Cancers Pancreas Cancer Cancer, Pancreas Cancers, Pancreas Pancreatic Cancer Cancer, Pancreatic Cancers, Pancreatic Pancreatic Cancers Cancer of the Pancreas OR "Pancreatic Neoplasm" [Title/Abstract] OR "Pancreatic Neoplasms" [Title/Abstract] OR "Prostatic Cancer" [Title/Abstract] OR "Prostatic Cancers" [Title/Abstract] OR "Pancreatic adenocarcinoma" [Title/Abstract]
Limiti: da gennaio 2009; humans		

3 EMBASE search strategy

"positron emission tomography"/syn OR "fluorodeoxyglucose f 18"/exp OR "fluorodeoxyglucose f 18"/syn OR "computer assisted emission tomography"/exp OR "computer assisted emission tomography" OR pet OR "pet scans" OR "pet scanner" OR	AND	"Pancreatic Neoplasms"/de, syn, Keyword OR "Pancreatic Neoplasms"/exp

“pet scan” OR “pet/ct scan” OR “pet/ct scans” OR “pet/ct” OR OR“positron emission tomography/computed tomography” OR OR pet NEAR/4 scan* OR pet NEAR/4 ct		OR “pancreatic cancer”: ti, ab. OR “pancreatic neoplasm”: ab:ti OR “pancreatic adenocarcinoma” : ab:ti OR “pancreatic neoplasms”: :ab:ti OR “pancreatic cancers”: :ab:ti OR Pancreatic NEAR/4 cancer* OR Pancreatic NEAR/4 neoplasm OR Pancreatic NEAR/4 cancers OR Pancreatic NEAR/4 neoplasms
Limiti: da gennaio 2009; humans “article” OR “review”/it OR “short survey”		

APPENDIX 10

FDG-PET/CT for staging of colorectal cancer

Characteristics of included studies

HTA report - ASSR colorectal cancer 2011

Document ID	ASSR-RER 2011 - Colon cancer
Objectives	to define criteria for appropriate use of FDG-PET for patients with colorectal cancer
Methods	A panel of experts working in Health Trusts and Teaching Hospitals of Emilia-Romagna was convened to discuss and agree on the methodology for a research programme aimed at defining the criteria for appropriate use of PET in colorectal cancer. On the basis of the clinical pathway of patients with colorectal cancer the panel examined and assessed the role of FDG-PET for 9 clinical indications (diagnosis of primary colorectal cancer, N staging of primary colorectal cancer, M staging of locally advanced colorectal cancer, target Volume definition of curative radiation treatment in patients with rectal cancer, during treatment evaluation of early response to therapy of liver metastases in colorectal cancer, end of treatment evaluation of response to neoadjuvant therapy for rectal cancer, evaluation of residual disease following ablative treatment of liver metastases, follow up in patients with no suspicion of recurrence, staging of suspect distant recurrence in patients treated for colorectal cancer). The following databases were searched for the period between January 2006 and September 2010: Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE - The Cochrane Library); Health Technology Assessment Database; Cochrane Central Register of Controlled Trials; National Library of Medicine's Medline database (PubMed); Elsevier's Embase. Language restrictions: English, Italian, French and Spanish. Selection criteria Type of studies: systematic reviews, RCTs, CCTs, cross-sectional diagnostic studies, prospective or retrospective cohort studies, case series of at least 10 patients Participants: patients with breast cancer Intervention: FDG-PET or CT/PET Reference standard: histology or clinical follow up Comparator: any other imaging technique Outcomes> sensitivity, specificity, LR, metabolic/tumor response, time to recurrence, local, local-regional and distant recurrence, disease free survival, disease survival, overall survival Assessment of methodological quality of studies The following criteria have been used for the quality assessment of different study designs. Systematic reviews: criteria drawn from the AMSTAR checklist Diagnostic cross sectional studies: criteria drawn from the QUADAS checklist

	Randomized controlled trials: criteria suggested by the Cochrane Handbook Case control studies and cohort studies: criteria drawn from the New Castle-Ottawa checklist Level of evidence for estimates of diagnostic accuracy were assigned according to GRADE categorization of the quality of evidence Each member of the panel voted the level of appropriateness for each clinical question. Two rounds of votes were requested for the judgment of appropriateness and results were analysed using the RAND/UCLA Appropriateness Method. The use of FDG-PET for a specific clinical indication was judged was judged as *appropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 7-9 score region as *inappropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 1-3 score region. Finally the use of FDG-PET was judged as *uncertain* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 4-6 score region or when no agreement was reached after the second round of voting. Clinical indications for which the panel does not reach an agreement on level of appropriateness after two rounds of voting also fall in the *uncertain* category.
Conclusions	N STAGING OF PRIMARY COLORECTAL CANCER - INAPPROPRIATE One systematic review and six primary studies evaluating FDG-PET's accuracy in N staging of primary colorectal cancer have been retrieved. However the panel established that there is no diagnostic role of FDG-PET in this clinical indication and unanimously agreed to judge its use as inappropriate. M STAGING OF LOCALLY ADVANCED COLORECTAL CANCER - APPROPRIATE After an initial strong disagreement, the panel agreed during the second meeting in rating the use of FDG-PET in staging patients with locally advanced primary colorectal cancer as appropriate. The disagreement was resolved through a clearer definition of the diagnostic role of FDG/PET for the selection of patients who would most benefit from radical surgery. The impact on survival obtained with appropriate surgical resection of localized disease and resectable metastases was in fact the only outcome considered critical (median score of 8; range 2-9), while remaining outcomes for true and false positives and for false negatives were judged important. The level of evidence for estimates of FDG-PET's diagnostic accuracy was moderate.
Notes	Meta-analysis of diagnostic accuracy estimates was not performed

SR - Brush 2011 – N staging

Disease	colorectal cancer
Index test	FDG-PET/CT
Comparators	none
Reference standard	histopathology following surgical resection and regional lymph node dissection
Target	diagnostic accuracy for N staging (regional lymph nodes) The systematic review assesses also diagnostic accuracy for distant metastases but only in patients with recurrent disease. Thus this question is not considered in this document.
Studies included	diagnostic accuracy studies with prospective or retrospective set-up of study
Years covered by the search	up to May 2009
Comprehensive bibliographic search: at least two databases searched	Yes (BIOSIS Previews; CINAHL Plus; The Cochrane Library; Compendex; ProQuest Dissertations and Theses; EMBASE; Global Health; Global Health Library regional indexes; Index to Theses; Inspec; MEDLINE; metaRegister of Current Controlled Trials; National Technical Information Services; OpenSIGLE; UK Clinical Research Network; Web of Science)
Characteristics of included studies clearly reported in tables	Yes
Methodological quality of primary studies assessed; criteria reported	Yes (QUADAS tool)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	N staging: 2 studies
Design of included studies	1 study retrospective design; 1 study unclear design
N. of included patients	141 (104 with rectal cancer, 37 with colon cancer)
Diagnostic accuracy results (with heterogeneity)	1 study FDG-PET/CT Proximal node staging

	Sensitivity 51.0% (95% CI 36.0-66.0%) Specificity 85.0% (95% CI 72.0-92.0%) Distal node staging Sensitivity 62.0% (95% CI 30.0-86.0%) Specificity 92.0% (95% CI 84.0-96.0%) 1 study FDG-PET/CT Sensitivity 85.0% (95% CI 69.0-93.0%) Specificity 42.0% (95% CI 23.0-67.0%) contrast-enhanced FDG-PET/CT Sensitivity 85.0% (95% CI 69.0-93.0%) Specificity 68.0% (95% CI 46.0-84.0%)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Unclear (1 study retrospective design, 1 study unclear design) Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Yes Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition?

		Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - N staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Mainenti 2011	FDG-PET/CT	34	patients with colorectal cancer scheduled for surgery	75.0	83.3

PS - Mainenti 2011

Clinical features and settings	colorectal cancer; Country: Italy
Participants	34 (20 men and 14 women; age range, 29-81 years; mean age: 63 years) with a histologically proven diagnosis of colorectal adenocarcinoma and scheduled for surgery. The regional distribution of the 37 tumors (two synchronous lesions were found in 3 out of 34 patients) was as follows: rectum (n = 6), rectosigmoid colon junction (n = 4), sigmoid colon (n = 15), descending colon (n = 3), transverse colon (n = 1), hepatic flexure (n = 3), ascending colon (n = 2) and caecum (n = 3). Five out of 37 (13.5%) tumors were classified as stage T1, 5 out of 37 (13.5%) as stage T2, 21 out of 37 (56.8%) as stage T3 and 6 out of 37 (16.2%) as stage T4. All three adenocarcinomas with a mucinous component were classified as T4. Twenty one out of 37 (57%) lesions were classified as N- and 16 out of 37 (43%) as N+ (13/16 as N1 and 3/16 as N2).
Study design	diagnostic cross sectional study with prospective recruitment; surgery was scheduled within 10 days of the examination, with the exception of three patients with rectal cancer who underwent neoadjuvant radio-chemotherapy after PET/CT and before surgery
Target condition and reference standard(s)	N staging (regional lymph nodes status) Reference standard: surgical findings and histopathological analysis of the surgical specimens
Index and comparator tests	FDG-PET/CT; comparators: endoscopic ultrasonography, thoracic and abdominal CT
Follow-up	not applicable for N staging; not reported for for extra regional lymph nodes staging
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT	Low risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Yes

SELECTION (risk of bias)		Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? not applicable Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review

		question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

Characteristics of excluded studies

Bamba 2011

Reason for exclusion	retrospective study

Chan 2011

Reason for exclusion	document of recommendations

Dirisamer 2010

Reason for exclusion	retrospective study

Eglinton 2010

Reason for exclusion	study of change in management

Floriani 2010

Reason for exclusion	all included studies with patients at recurrence phase

Hunter 2011

Reason for exclusion	study on risk of metastases

Kim 2011

Reason for exclusion	retrospective study

Kochhar 2010

Reason for exclusion	retrospective study

Mainenti 2010

Reason for exclusion	Included in ASSR-RER 2011 HTA report

Niekel 2010

Reason for exclusion	unclear condition of patients in included studies (staging or recurrence or both)

Nozawa 2012

Reason for exclusion	full text cannot be found

Patel 2011

Reason for exclusion	all included studies with patients at recurrence phase

Ramos 2011

Reason for exclusion	per-lesion analysis

Strasberg 2010

Reason for exclusion	narrative review

Van der Pas 2011

Reason for exclusion	another clinical question

Wiering 2010

Reason for exclusion	consider only FDG-PET

Yu 2012

Reason for exclusion	per-lesion analysis

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”:ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw or
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. “Colorectal Neoplasms” [Mesh descriptor NoExp]
10. “Colonic Neoplasms” [Mesh descriptor NoExp]
11. “Rectal Neoplasms” [Mesh descriptor explode all trees]
12. 9/12 OR
13. 8 AND 12

Publication date: January 2010 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18”[Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg pet”[All Fields]
17. “Positron-Emission Tomography”[Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. “colorectal carcinoma”[Title/Abstract]
26. “colorectal neoplasm”[Title/Abstract]
27. “colorectal neoplasms”[Title/Abstract]
28. “colorectal cancer”[Title/Abstract]
29. “colorectal cancers”[Title/Abstract]
30. “colonic neoplasm”[Title/Abstract]
31. “colonic neoplasms”[Title/Abstract]
32. “cancer of colon”[Title/Abstract]
33. “colon cancers”[Title/Abstract]
34. “colon cancer”[Title/Abstract]
35. “sigmoid neoplasm”[Title/Abstract]
36. “sigmoid neoplasms”[Title/Abstract]
37. “sigmoid cancer”[Title/Abstract]
38. “sigmoid cancers”[Title/Abstract]
39. “sigmoidal cancer”[Title/Abstract]
40. “cancer of sigmoid”[Title/Abstract]

41. “rectal neoplasm”[Title/Abstract]
42 “rectal neoplasms”[Title/Abstract]
43. “rectal cancer”[Title/Abstract]
44. “rectal cancers”[Title/Abstract]
45. “rectum cancer”[Title/Abstract]
46. “rectum cancers”[Title/Abstract]
47. “cancer of rectum”[Title/Abstract]
48. “Sigmoid Neoplasms”[Mesh:noexp]
49. “Colorectal Neoplasms”[Mesh:noexp]
50. “Colonic Neoplasms”[Mesh:noexp]
51. “Rectal Neoplasms”[Mesh]
52. 25/51 OR
53. 24 AND 25
54. “editorial”[Publication Type]
55. “comment”[Publication Type]
56. “letter”[Publication Type]
57. 54/56 OR
58. 53 NOT 57

Limits: humans
Publication date: January 2010 - March 2012
Languages: English, French, Italian, Spanish

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “fluorodeoxyglucose f 18”/exp
3. “fluorodeoxyglucose f 18”/syn
4. “computer assisted emission tomography”/exp
5. “computer assisted emission tomography” OR
6. pet
7. “pet scans”
8. “pet scanner”
9. “pet scan”
10. “pet/ct scan”
11. “pet/ct scans”
12. “pet/ct”
13. “positron emission tomography/computed tomography”
14. pet NEAR/4 scan*
15. pet NEAR/4 ct
16. 1/15 OR
17. “colon cancer”/de, not exp
18. “colon adenocarcinoma”/de, not exp
19. “colon carcinogenesis”/de, not exp
20. “colon carcinoma”/de, not exp
21. “colorectal cancer”/de, not exp
22. “colorectal carcinoma”/de, not exp
23. “sigmoid carcinoma”/de, not exp
24. “cecum cancer”/de, exp
25. “rectum carcinoma”/de, exp
26. “anus cancer”/ de, exp
27. “colon cancer”:ab,ti
28. “colon adenocarcinoma”:ab,ti
29. “colon carcinogenesis”:ab,ti
30. “colon carcinoma”:ab,ti
31. “colorectal cancer”:ab,ti
32. “colorectal carcinoma”:ab,ti
33. “sigmoid carcinoma”:ab,ti
34. “cecum cancer”:ab,ti
35. “rectum carcinoma”:ab,ti
36. “anus cancer”:ab,ti
37. “colonic cancer”:ab,ti
38. “rectosigmoid adenocarcinoma”:ab,ti

39. “carcinoma coli”:ab,ti
40. “anal cancer”:ab,ti
41. “caecal cancer”:ab,ti
42. “caecum cancer”:ab,ti
43. “cecum sarcoma”:ab,ti
44. 17/43 OR
45. 16 AND 44
46. 45 AND (“article” OR “review” OR “short survey” OR “in press article”)

Limits: Humans
Publication date: January 2010 - March 2012
Languages: English, French, Italian, Spanis

APPENDIX 11

FDG-PET/CT for staging of renal cancer

HTA report - KCE 2009

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients; • Absence of adequate reference standard; • Absence of patient-based analysis;

	• Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	The 2009 KCE report conclusions for kidney cancer staging are based on the AHRQ 2008 report (AHRQ 2008; KCE 2009). 2009 KCE report concluded that the evidences on staging are limited to small studies of low quality. For the initial staging, PET can be useful when CT and/or bone scan are dubious, although this results should be confirmed by prospective trials.
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

Characteristics of excluded studies

Ansquer 2010

Reason for exclusion	No kidney cancer

Lodde 2010

Reason for exclusion	No kidney cancer

Ozulker 2011

Reason for exclusion	No staging

Ye 2010

Reason for exclusion	No primary tumor

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

Kidney Tumor: search strategy CDSR, DARE, HTA database, CENTRAL		
“Positron-Emission Tomography” [MeSH descriptor explode all trees] OR “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees] OR. “positron emission tomography”: ti,ab,kw OR. pet*: ti,ab,kw OR. pet scan*: ti,ab,kw OR. “Fluorodeoxyglucose F18”: ti,ab,kw OR. fdg NEAR/2 18: ti,ab,kw	AND	Kidney Neoplasms [Mesh explodes all trees] Kidney Neoplasm Neoplasm, Kidney Renal Neoplasms Neoplasm, Renal Neoplasms, Renal Renal Neoplasm Neoplasms, Kidney Cancer of Kidney Kidney Cancers Renal Cancer Cancer, Renal Cancers, Renal Renal Cancers Cancer of the Kidney Kidney Cancer Cancer, Kidney Cancers, Kidney OR Kidney Neoplasm*: ti,ab,kw OR Kidney Cancer*: ti,ab,kw OR Kidney Tumor*:ti,ab,kw OR Kidney carcinoma*: ti,ab,kw OR Renal Neoplasm*:ti,ab,kw OR Renal Cancer*:ti,ab,kw OR Renal Tumor*:ti,ab,kw OR Renal Carcinoma*:ti,ab,kw

		OR Kidney NEAR/4 cancer*: ti,ab,kw OR Kidney NEAR/4 neoplasm*: ti,ab,kw OR Renal NEAR/4 cancer*: ti,ab,kw OR Renal NEAR/4 neoplasm*: ti,ab,kw OR Kidney NEAR/4 tumor*: ti,ab,kw OR Renal NEAR/4 tumor*: ti,ab,kw OR Kidney carcinoma*:ti,ab,kw OR Renal carcinoma*:ti,ab,kw

From genuary 2009

2 MEDLINE search strategy

Search strategy kidney cancer /Medline		
"Fluorodeoxyglucose F18"[Mesh] **OR** "2-Fluoro-2-deoxyglucose" [All Fields] **OR** "18F Fluorodeoxyglucose" [All Fields] **OR** "F 18 Fluorodeoxyglucose" [All Fields] **OR** Fludeoxyglucose* [All Fields] **OR** "2 fluoro 2 deoxy d glucose"[All Fields] **OR** 18fluorodesoxyglucose*[All Fields] **OR** fluorodeoxyglucose*[All Fields] **OR** "fluorine 18 fluorodeoxyglucose" [All Fields] **OR** 18f dg*[All Fields]) **OR** 18fluorodeoxyglucose*[All Fields] **OR** 18fdg [All Fields] **OR** 18 fdg* [All Fields] **OR** fdg 18* [All Fields] **OR** fdg/* [All Fields] **OR** "fdg pet"[All Fields] **OR** "Positron-Emission Tomography"[Mesh] **OR**	AND	Kidney Neoplasms [Mesh **explodes all trees]** **contiene** Kidney Neoplasm Neoplasm, Kidney Renal Neoplasms Neoplasm, Renal Neoplasms, Renal Renal Neoplasm Neoplasms, Kidney Cancer of Kidney Kidney Cancers Renal Cancer Cancer, Renal Cancers, Renal Renal Cancers Cancer of the Kidney Kidney Cancer Cancer, Kidney

“positron emission tomography” [title/abstract] **OR** pet [title/abstract] **OR** “pet scan” [All Fields] **OR** “pet scans” [All Fields] **OR** “pet scanner” [All Fields] **OR** petscan [All Fields]		Cancers, Kidney OR “Kidney Neoplasm” [Title/Abstract] OR “Kidney Neoplasms” [Title/Abstract] OR “Renal Neoplasm” [Title/Abstract] OR “Renal Neoplasms” [Title/Abstract] OR “Kidney Cancer” [Title/Abstract] OR “Kidney Cancers” [Title/Abstract] OR “Renal Cancer” [Title/Abstract] OR “Renal Cancers” [Title/Abstract] OR “Kidney tumor” [Title/Abstract] OR “Renal tumor” [Title/Abstract] OR “Kidney tumors” [Title/Abstract] OR “Renal tumors” [Title/Abstract] OR “Kidney carcinoma” [Title/Abstract] OR “Renal carcinoma” [Title/Abstract]

3 EMBASE search strategy

Search strategy kidney cancer / Embase		
“positron emission tomography”/syn OR “fluorodeoxyglucose f 18”/exp OR “fluorodeoxyglucose f 18”/syn OR “computer assisted emission tomography”/exp OR “computer assisted emission tomography” OR pet OR “pet scans” OR “pet scanner” OR “pet scan” OR “pet/ct scan” OR “pet/ct scans” OR “pet/ct” OR OR“positron emission tomography/computed tomography” OR OR pet NEAR/4 scan* OR pet NEAR/4 ct	AND	“Kidney Neoplasms”/exp OR “Kidney cancer”/de, syn, Keyword” OR “Kidney cancers”/de, syn, Keyword” OR “Kidney Neoplasms”/de, syn, Keyword OR “Kidney neoplasm”/de,syn, keyword OR “Kidney tumor“/de,syn, keyword OR “Kidney tumors”/de,syn, keyword OR “Kidney carcinoma” ”/de,syn, keyword OR “Renal cancer”/de, syn, Keyword OR

		“Renal cancers”/de,syn;keyword OR “Renal tumor ”/de,syn, keyword OR “Renal tumors ”/de,syn, keyword OR “Renal neoplasm”/de,syn, keyword OR “Renal neoplasms”/de,syn, keyword OR “Kidney neoplasm”: ab:ti OR “Kidney neoplasms”: ab:ti OR “Kidney cancer”: ab:ti; OR “Kidney cancers”: ab:ti OR “Kidney tumor”: ab:ti; OR “Kidney tumors”: ab:ti, OR “Kidney carcinoma”: ab:ti OR “Kidney carcinomas” ”: ab:ti OR “Renal neoplasms”: ab:ti OR “Renal neoplasm” : ab:ti OR “Renal cancers”: :ab:ti OR “Renal cancer”: :ab:ti OR “Renal tumors” :ab:ti OR “Renal tumor”: :ab:ti OR “Renal carcinoma”:ab:ti OR “Renal carcinomas”:ab:ti Kidney NEAR/4 neoplasm* OR

		Kidney NEAR/4 cancer* OR Kidney NEAR/4 tumor OR Renal NEAR/4 neoplasm* OR Renal NEAR/4 cancer* OR Renal NEAR/4 tumor*

APPENDIX 12

FDG-PET/CT for staging of bladder cancer

Characteristics of included studies

1 HTA report - KCE 2009 bladder cancer

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients;

	• Absence of adequate reference standard; • Absence of patient-based analysis; • Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	**Staging:** the AHRQ 2008 report identified 2 prospective studies on the use of PET for the primary staging of bladder cancer. Sensitivity was 53% and 77%, specificity was 72% and 94%. KCE 2009 authors identified no additional primary studies and concluded that the evidence on the use of PET/CT is too limited to base recommendations on. **Clinical effectiveness:** authors reported no evidence for the use of PET and PET/CT (KCE 2009).
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

PS - N staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Swinnen 2009	FDG-PET/CT	51	patients with histologically proven transitional cell carcinoma (TCC; T2 or higher) or recurrent high-risk superficial TCC (T1G3 with or without Tis) of the bladder.	46.2 (95% CI 22.4–71.3)	97.4 (95% CI 88.1–99.9)
	CT			46.2 (95% CI 22.4–71.3)	92.1 (95% CI: 80.9–97.8)

PS - Swinnen 2009

Clinical features and settings	Bladder carcinoma: Belgium
Participants	51 patients (male: 43) with histologically proven transitional cell carcinoma (TCC; T2 or higher) or recurrent high-risk superficial TCC (T1G3 with or without Tis) of the bladder during the period from April 2004 until December 2007 were included; mean age (range): 66 (48-82); All patients underwent cystectomy and an extended lymphadenectomy with eight separate nodal sites was labelled as paraaortic; paracaval; pelvic (right and left), including external and internal iliac; obturator fossa (right and left); and presacral. The eighth region was the perivesical tissue attached to the cystectomy specimen. pT1 pN0 (n):12 pT2 pN0 (n): 20 pT3 pN0 (n): 6 pT4 pN0 (n): 0 pT1 pN+ (n): 0 pT2 pN+ (n): 2 pT3 pN+ (n): 7 pT4 pN+ (n): 4 pT: pathologic classification of the primary tumor status pN0: pathologic lymph node–negative pN+: pathologic lymph node–positive
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	paraaortic; paracaval; pelvic (right and left), including external and internal iliac; obturator fossa (right and left); and presacral lymph node metastases; Reference standard: pathological proof and follow-up.
Index and comparator tests	FDG-PET/CT; comparator: CT.

Follow-up	One patient had 25 months follow up; for the remaining patients the follow up is not stated.
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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** Could the selection of patients have introduced bias? **RISK:** LOW=YES/HIGH=NO/**UNCLEAR**
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? **Unclear** If a threshold was used, was it pre-specified? **Unclear** Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
3A. REFERENCE STANDARD (risk of bias)	Low risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? **Yes** Were the reference standard results interpreted without knowledge of the results of the index test? **Unclear** Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? **Unclear** Did all patients receive a reference standard?

		Yes Did patients receive the same reference standard? **Yes** Were all patients included in the analysis? **Yes** Could the patient flow have introduced bias? RISK: **LOW=YES**/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? **CONCERN: LOW=YES**/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: **LOW=YES**/HIGH=NO**/**UNCLEAR

SR - Lu 2011 - Any staging

Disease	bladder carcinoma
Index test	FDG-PET and/or FDG-PET/CT
Comparators	not reported
Reference standard	pathology (from biopsy or surgery), or follow-up
Target	lymph node and distant metastases (N - M staging)
Studies included	6 diagnostic studies with prospective or retrospective patients recruitment to perform accuracy diagnostic test metanalysis for recurrence and/or staging/restaging (three studies for staging**:** PS - Apolo 2010, PS - Kibel 2009 with FDG-PET/CT as index test and one study: PS - Drieskens 2005 with FDG-PET as index test).
Years covered by the search	up to July 2011 (submission date)
Comprehensive bibliographic search: at least two databases searched	PubMed/MEDLINE and EBM Review
Characteristics of included studies clearly reported in tables	YES
Methodological quality of primary studies assessed; criteria reported	YES Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests: Recommended Methods. http://www.cochrane.org/cochrane/sadtdoc1.htm (accessed on June 6, 1996).
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	YES
N. of included studies	6 primary diagnostic studies; Two studies included in diagnostic accuracy of FDG PET and PET/CT metanalysis for detection of primary tumor; 5 studies included in diagnostic accuracy of FDG PET and PET/CT metanalysis for staging or restaging (lymph node and distant metastases).
Design of included studies	diagnostic accuracy studies with prospective (PS - Apolo 2010; PS - Kibel 2009; PS - Drieskens 2005) or retrospective patients recruitment.
N. of included patients	236 patients (219 in staging/restaging meta-analysis)
Diagnostic accuracy	**N - M Staging** [diagnostic accuracy of FDG PET and PET/CT metanalysis for

results (with heterogeneity)	staging or restaging (lymph node and distant metastases) - Five studies). **Pooled sensitivity** : 0.82 (95% CI: 0.72–0.89) I^2:79.6% **Pooled specificity** : 0.89 (95% CI: 0.81–0.95) I^2:65.6% **Diagnostic test accuracy (sensitivity and specificity) (three studies for staging).** PS - Apolo 2010 sensitivity : 0.81 (95% CI: 0.63–0.93) specificity : 0.94 (95% CI: 0.70–1.00) PS - Kibel 2009 sensitivity : 0.70 (95% CI: 0.35–0.93) specificity : 0.94 (95% CI: 0.79–0.99) PS - Drieskens 2005 sensitivity : 0.53 (95% CI: 0.27–0.79) specificity : 0.72 (95% CI: 0.51–0.88)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	High risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? **Unclear** Was a case-control design avoided? **No** Did the study avoid inappropriate exclusions? **Unclear** Could the selection of patients have introduced bias? RISK: LOW=YES/**HIGH=NO**/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? **Yes** If a threshold was used, was it pre-specified? **Unclear** Could the conduct or interpretation of the index test have introduced bias? RISK: **LOW=YES**/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	High risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target

		condition? **Yes** Were the reference standard results interpreted without knowledge of the results of the index test? **No** Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/**HIGH=NO**/UNCLEAR
4. FLOW AND TIMING (risk of bias)	High risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? **Unclear** Did all patients receive a reference standard? **No** Did patients receive the same reference standard? **Unclear** Were all patients included in the analysis? **No** Could the patient flow have introduced bias? RISK: LOW=YES/**HIGH=NO/**UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Unclear risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/U**NCLEAR**
2.B INDEX TEST(S) (concern of applicability)	Unclear risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**
3B. REFERENCE STANDARD (concern of applicability)	Unclear risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**

Characteristics of excluded studies

Almuhaideb 2011

Reason for exclusion	FDG PET/CT imaging overview in oncology

Bostrom 2010

Reason for exclusion	Staging technics overview

Boujelbene 2011

Reason for exclusion	Narrative review

Jensen 2011

Reason for exclusion	Retrospective study

Lodde 2010

Reason for exclusion	Mixed population: participants enrolled for staging of primary cancer less than 80%.

Moses 2011

Reason for exclusion	Imaging technics overview

Rioja 2010

Reason for exclusion	Narrative review

Zouhair 2010

Reason for exclusion	Narrative review

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”: ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. Urinary Bladder [Mesh explodes all trees]
10. “urinary bladder”:ti,ab,kw
11. Bladder NEAR/4 cancer*: ti,ab,kw
12. Bladder NEAR/4 neoplasm*: ti,ab,kw
13. 9/12 OR
14. Neoplasm*: ti,ab,kw
15. Cancer*:ti,ab,kw
16. Carcinoma:ti,ab,kw
17. 14/16 OR
18. 8 AND 13 AND 17

Publication date: January 2009 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18” [Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg-pet”[All Fields]
17. “Positron-Emission Tomography” [Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. Urinary Bladder [Mesh explodes all trees]
26. “Bladder neoplasm” [Title/Abstract]
27. “Bladder Neoplasms” [Title/Abstract]
28. “Bladder Cancer” [Title/Abstract]
29. “Bladder Cancers” [Title/Abstract]
30. 25/29 OR
31. 24 AND 30

Limit: Humans

Publication date: January 2009 - March 2012

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “positron emission tomography”/exp
3. “fluorodeoxyglucose f 18”/exp
4. “fluorodeoxyglucose f 18”/syn
5. “computer assisted emission tomography”/exp
6. “computer assisted emission tomography”/tw
7. pet/tw
8. “pet scans”/tw
9. “pet scanner”/tw
10. “pet scan”/tw
11. “pet/ct scan”/tw
12. “pet/ct scans”/tw
13. “pet/ct”/tw
14. “positron emission tomography/computed tomography”/tw
15. pet NEAR/4 scan*
16. pet NEAR/4 ct
17. 1/16 OR
18. “bladder neoplasms”/de, syn, Keyword
19. “bladder Neoplasms”/exp
20. “bladder cancer”/de, syn, Keyword”
21. “bladder cancers”/de, syn, Keyword”
22. “bladder Neoplasms”/de, syn, Keyword
23. “bladder neoplasm”/de, syn, Keyword
24. “bladder cancer”: ti, ab.
25. “bladder neoplasm”: ab:ti
26. “bladder neoplasms”: ab:ti
27. “bladder cancers”: ab:ti
28. Bladder NEAR/4 cancer*
29. Bladder NEAR/4 neoplasm*
30. 18/29 OR
31. 17 AND 30

Limit: Humans

Limit: “article” OR “review”/it OR “short survey”

Publication date: January 2009 - March 2012

APPENDIX 13

FDG-PET/CT for staging of uterine cancer

Characteristics of included studies

HTA report - KCE 2009 uterine cancer

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients;

	• Absence of adequate reference standard; • Absence of patient-based analysis; • Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	KCE authors identified two primary studies. One small retrospective study (Torizuka 2006) compared PET with MRI for the assessment of myometrial infiltration in patients with clinical stage I uterine corpus cancer. Histopathology was used as reference standard. The study possibly suffered from selection bias. Sensitivity for PET and MRI was 83% and 100% respectively, while specificity was 88% and 69%. All 95% confidence intervals were overlapping. One prospective study (Kitajima 2008) evaluated the use of PET/CT for the N-staging of patients with primary endometrial cancer. Histopathology was used as reference standard. Sensitivity and specificity were found to be 50% and 87% respectively. The HTA document concluded that **Staging**: the evidence on the use of PET and PET/CT is too limited to base recommendations on. **Clinical effectiveness:** authors reported no evidence for the use of PET and PET/CT.
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

SR - Chang 2012 - N staging

Disease	uterine (endometrial) cancer
Index test	FDG-PET and FDG-PET/CT
Comparators	not reported
Reference standard	histopathological proof and/or clinical follow-up
Target	lymph node (pelvic and/or paraaortic) metastases (N staging)
Studies included	7 diagnostic studies with prospective or retrospective patients recruitment to perform accuracy diagnostic test metanalysis for N staging (two studies underwent as index test FDG-PET alone**:** Inubashiri 2009, Horowitz 2004; four studies FDG-PET/CT: PS - Signorelli 2009; Nakamura 2010; Kitajima 2008; Picchio 2010) and one study: Suga 2011, FDG-PET or FDG-PET/CT).
Years covered by the search	From January 1998 to March 2011.
Comprehensive bibliographic search: at least two databases searched	No; only MEDLINE database
Characteristics of included studies clearly reported in tables	YES
Methodological quality of primary studies assessed; criteria reported	YES modified QUADAS tool according to KCE 2009: twelve methodological quality items were assessed for each study using the scores “yes,” “no,” or “unclear” for each item. “No” and “unclear” responses were interpreted as having not achieved the quality item. A quality score for each study was expressed as a percentage of the maximum score of 12.
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	YES
N. of included studies	7 primary accuracy diagnostic studies
Design of included studies	diagnostic accuracy studies with prospective (four studies: Inubashiri 2009; PS - Signorelli 2009; Kitajima 2008; Horowitz 2004), retrospective (Suga 2011; Picchio 2010) or not reported (one study: Nakamura 2010) patients recruitment.
N. of included patients	243 patients

Diagnostic accuracy results (with heterogeneity)	**N Staging** [diagnostic accuracy of FDG PET and PET/CT metanalysis for staging (lymph node metastases) - Seven studies]. **Pooled sensitivity** : 63.0% (95% CI, 48.7–75.7%) I^2 :48.3% (p=0.071) **Pooled specificity** : 94.7% (95% CI, 90.4–97.4%) $I^{2:}$ 45.7% (p=0.087)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? **Unclear** Was a case-control design avoided? **Unclear** Did the study avoid inappropriate exclusions? **Unclear** Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? **Yes** If a threshold was used, was it pre-specified? **Unclear** Could the conduct or interpretation of the index test have introduced bias? RISK: **LOW=YES**/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? **Yes** Were the reference standard results interpreted without knowledge of the results of the index test? **Unclear** Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
4. FLOW AND TIMING (risk of bias)	High risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and

		reference standard? **Unclear** Did all patients receive a reference standard? **Unclear** Did patients receive the same reference standard? **No** Were all patients included in the analysis? **No** Could the patient flow have introduced bias? RISK: LOW=YES/**HIGH=NO/**UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Unclear risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR

Characteristics of excluded studies

Alt 2011

Reason for exclusion	narrative review

Basu 2009

Reason for exclusion	narrative review

Basu 2010

Reason for exclusion	narrative review

Brocker 2011

Reason for exclusion	narrative review

Brooks 2009

Reason for exclusion	narrative review

Caroli 2010

Reason for exclusion	narrative review

Kitajima 2009

Reason for exclusion	mixed population (30 patients with endometrial cancer and 15 with cervical cancer)

Kitajima 2010

Reason for exclusion	narrative review

Kitajima 2011a

Reason for exclusion	retrospective study

Kitajima 2011b

Reason for exclusion	narrative review

Klumpp 2012

Reason for exclusion	target condition is not N or M staging but extent of peritoneal carcinomatosis

Lee HJ 2011

Reason for exclusion	case-control study

Lee JH 2011

Reason for exclusion	guidelines

Ma 2011

Reason for exclusion	narrative review

Rockall 2012

Reason for exclusion	narrative review

Sohaib 2010

Reason for exclusion	narrative review

Tsujikawa 2011

Reason for exclusion	narrative review

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”: ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. Uterine Neoplasms
10. Uterine Neoplasm*: ti,ab,kw
11. Uterine Cancer*: ti,ab,kw
12. uterine Tumor*:ti,ab,kw
13. uterine carcinoma*: ti,ab,kw
14. uterus Neoplasm*:ti,ab,kw
15. Uterus Cancer*:ti,ab,kw
16. uterus Tumor*:ti,ab,kw
17. uterine NEAR/4 cancer*: ti,ab,kw
18. uterine NEAR/4 neoplasm*: ti,ab,kw
19. uterus NEAR/4 cancer*: ti,ab,kw
20. uterus NEAR/4 neoplasm*: ti,ab,kw
21. uterine NEAR/4 tumor*: ti,ab,kw
22. uterine NEAR/4 tumor*: ti,ab,kw
23. Gynecological Tumor*:ti,ab,kw
24. 8 AND 23

Publication date: January 2009 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18” [Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg-pet”[All Fields]
17. “Positron-Emission Tomography” [Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. Uterine Neoplasms [Mesh explodes all trees]
26. “Uterine Neoplasm” [Title/Abstract]
27. “Uterine Neoplasms” [Title/Abstract]

28. “Uterus Neoplasm” [Title/Abstract]
29. “Uterus Neoplasms” [Title/Abstract]
30. “Uterine Cancer” [Title/Abstract]
31. “Uterine Cancers” [Title/Abstract]
32. “Uterus Cancer” [Title/Abstract]
33. “Uterus Cancers” [Title/Abstract]
34. “uterine tumor” [Title/Abstract]
35. “uterus tumor” [Title/Abstract]
36. “gynecological tumor” [Title/Abstract]
37. “gynecological tumors” [Title/Abstract]
38. 25/37 OR
39. 24 AND 38

Limit: Humans

Publication date: January 2009 - March 2012

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “positron emission tomography”/exp
3. “fluorodeoxyglucose f 18”/exp
4. “fluorodeoxyglucose f 18”/syn
5. “computer assisted emission tomography”/exp
6. “computer assisted emission tomography”/tw
7. pet/tw
8. “pet scans”/tw
9. “pet scanner”/tw
10. “pet scan”/tw
11. “pet/ct scan”/tw
12. “pet/ct scans”/tw
13. “pet/ct”/tw
14. “positron emission tomography/computed tomography”/tw
15. pet NEAR/4 scan*
16. pet NEAR/4 ct
17. 1/16 OR
18. “Uterine Neoplasms”/de, syn, Keyword
19. “Uterine Neoplasms”/exp
20. “uterine cancer”/de, syn, Keyword”
21. “uterine cancers”/de, syn, Keyword”
22. “uterus Neoplasms”/de, syn, Keyword
23. “uterus cancer”/de, syn, Keyword
24. “uterus cancers”/de,syn;keyword
25. “uterus neoplasm”/de,syn, keyword
26. “uterine tumor“/de,syn, keyword
27. “uterine tumors”/de,syn, keyword
28. “uterine cancer”: ti, ab.
29. “uterine neoplasm”: ab:ti
30. “uterine neoplasms” : ab:ti
31. “uterine cancers”: :ab:ti
32. “uterine tumor”: :ab:ti
33. “uterine tumors” :ab:ti
34. “uterus cancers”: :ab:ti
35. “uterus cancer”:ab:ti
36. “uterus neoplasm”:ab:ti
37. “uterus neoplasms”:ab:ti
38. “uterus tumor”: ab:ti
39. “uterus tumors”:ab:ti
40. “gynecological tumors”:ab:ti
41. uterine NEAR/4 neoplasm*
42. uterine NEAR/4 cancer*
43. uterus NEAR/4 neoplasm*

44. uterus NEAR/4 cancer*
45. uterus NEAR/4 tumor*
46. uterus NEAR/4 tumor*
47. 18/46 OR
48. 17 AND 47

Limit: Humans

Limit: “article” OR “review”/it OR “short survey”

Publication date: January 2009 - March 2012

APPENDIX 14

FDG-PET/CT for staging of cervical cancer

Characteristics of included studies

HTA report - KCE 2009 cervical cancer

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients; • Absence of adequate reference standard;

	• Absence of patient-based analysis; • Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	The KCE report included an AHRQ report published in 2008 (12 studies) and two systematic reviews of mixed quality. The AHRQ 2008 document identified 6 primary studies (published between 2003 and March 2008) on the use of PET/CT for the initial staging of cervical cancer. Two studies provided a patient-based analysis for N-staging: sensitivity was 50% and 100%, while specificity was 83% and 99%. In 3 studies involving 211 patients, sensitivity ranged from 60 to 100% and specificity was 94% in 2 studies (the third study reported no data) for the detection of extra-cervical and/or metastatic disease. KCE failed to identify new studies since the AHRQ report and concluded that a number of studies "reported a low sensitivity for pelvic lymph node staging, but a moderate sensitivity for extrapelvic lymph node staging. Specificity was consistently good across both lymph node regions (level 2). A good-quality systematic review found sentinel-node biopsy to be the most accurate technique for early-stage disease (level 2)" It is unclear however whether these findings refer to PET or PET/CT.
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

PS - N staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Signorelli 2012	FDG-PET/CT	159	Women with Ib1–IIa < 4 cm cervical carcinoma	32.1	96.9

PS - Signorelli 2011

Clinical features and settings	Signorelli 2011
Participants	159 women (median age 49 years) with Ib1–IIa < 4 cm cervical carcinoma
Study design	Prospective single centre consecutively recruited cohort from the San Raffaele Hospital in Milan, Italy
Target condition and reference standard(s)	Cervical cancer. Reference standard was intra operative histology performed by an operator blinded to imaging results
Index and comparator tests	18-FDG-PET/CT vs MRI for initial staging
Follow-up	Not mentioned
Notes	The authors conclude that "low sensitivity of 18F-FDG-PET/CT scan in depicting nodal metastases. 18F-FDG-PET/CT demonstrated a minimal clinical impact in the treatment planning and should not be incorporated as a routine imaging technique in the pre-treatmentmanagement of women with early stage cervical cancer"

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of patient selection: ϖ Was a consecutive or random sample of patients enrolled? Yes/No/Unclear ϖ Was a case-control design avoided? Yes/No/Unclear ϖ Did the study avoid inappropriate exclusions? Yes/No/Unclear Could the selection of patients have introduced bias? RISK: LOW
2A. INDEX TEST(S) (risk of	Low risk	Describe the index test and how it was conducted and

bias)		interpreted: ϖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear ϖ If a threshold was used, was it pre-specified? Yes/No/Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW
3A. REFERENCE STANDARD (risk of bias)	Low risk	Describe the reference standard and how it was conducted and interpreted: ϖ Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear ϖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: ϖ Was there an appropriate interval between index test(s) and reference standard? Yes/No/Unclear ϖ Did all patients receive a reference standard? Yes/No/Unclear ϖ Did patients receive the same reference standard? Yes/No/Unclear ϖ Were all patients included in the analysis? Yes/No/Unclear NO MENTION OF FOLLOW UP IN THE TEXT Could the patient flow have introduced bias? RISK: UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question?

		CONCERN: LOW
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: Unclear
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: Unclear

Footnotes

Characteristics of excluded studies

Ferrandina 2012

Reason for exclusion	Re-staging study

Kang 2010

Reason for exclusion	Searches overlap KCE searches

Kitajima 2009

Reason for exclusion	Mixed cancer study with no breakdown of data by pathology

Lee 2011

Reason for exclusion	Re-staging study

Leseur 2011

Reason for exclusion	Study on PET, in French

Olsen 2011

Reason for exclusion	Treatment study with biomarker comparison

Ozcan 2011

Reason for exclusion	Retrospective study

Rudmik 2011

Reason for exclusion	Diagnosis only study

Small 2010

Reason for exclusion	Not PET/CT study

Tatsumi 2009

Reason for exclusion	Retrospective study

Tsai 2010

Reason for exclusion	Treatment study

Ylmaz 2010

Reason for exclusion	Retrospective study

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

“Positron-Emission Tomography” [MeSH descriptor explode all trees] OR “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees] OR. “positron emission tomography”: ti,ab,kw OR. pet*: ti,ab,kw OR. pet scan*: ti,ab,kw OR. “Fluorodeoxyglucose F18”: ti,ab,kw OR. fdg NEAR/2 18: ti,ab,kw	AND	"Uterine Cervical Neoplasms"[Mesh]: contiene (Cervical Neoplasm, Uterine Cervical Neoplasms, Uterine Neoplasm, Uterine Cervical Neoplasms, Uterine Cervical Uterine Cervical Neoplasm Neoplasms, Cervical Cervical Neoplasms Cervical Neoplasm Neoplasm, Cervical Neoplasms, Cervix Cervix Neoplasms Cervix Neoplasm Neoplasm, Cervix Cancer of the Uterine Cervix Cancer of the Cervix Uterine **Cervical Cancer** Cancer, Uterine Cervical Cancers, Uterine Cervical **Cervical Cancer**, Uterine Cervical Cancers, Uterine Uterine Cervical Cancers Cancer of Cervix Cervix Cancer Cancer, Cervix Cancers, Cervix) OR Cervical Neoplasm*: ti,ab,kw OR Uterine Cervical : ti,ab,kw OR Cervix Neoplasm* : ti,ab,kw OR “Cancer of the Uterine Cervix” : ti,ab,kw OR “Cancer of the Cervix” : ti,ab,kw OR

		“Cervical Cancer*”: ti,ab,kw OR “Cancer of Cervix”: ti,ab,kw OR “Cervix Cancer”: ti,ab,kw

2 MEDLINE search strategy

“Fluorodeoxyglucose F18”[Mesh] **OR** “2-Fluoro-2-deoxyglucose” [All Fields] **OR** “18F Fluorodeoxyglucose” [All Fields] **OR** “F 18 Fluorodeoxyglucose” [All Fields] **OR** Fludeoxyglucose* [All Fields] **OR** “2 fluoro 2 deoxy d glucose”[All Fields] **OR** 18fluorodesoxyglucose*[All Fields] **OR** fluorodeoxyglucose*[All Fields] **OR** “fluorine 18 fluorodeoxyglucose” [All Fields] **OR** 18f dg*[All Fields]) **OR** 18fluorodeoxyglucose*[All Fields] **OR** 18fdg [All Fields] **OR** 18 fdg* [All Fields] **OR** fdg 18* [All Fields] **OR** fdg/* [All Fields] **OR** “fdg pet”[All Fields] **OR** “Positron-Emission Tomography”[Mesh] **OR** “positron emission tomography” [title/abstract] **OR** pet [title/abstract] **OR** “pet scan” [All Fields] **OR** “pet scans” [All Fields] **OR** “pet scanner” [All Fields] **OR** petscan [All Fields]	AND	“Uterine Cervical neoplasm ”[Mesh explodes all trees] contiene: (Cervical Neoplasm, Uterine Cervical Neoplasms, Uterine Neoplasm, Uterine Cervical Neoplasms, Uterine Cervical Uterine Cervical Neoplasm Neoplasms, Cervical Cervical Neoplasms Cervical Neoplasm Neoplasm, Cervical Neoplasms, Cervix Cervix Neoplasms Cervix Neoplasm Neoplasm, Cervix Cancer of the Uterine Cervix Cancer of the Cervix Uterine **Cervical Cancer** Cancer, Uterine Cervical Cancers, Uterine Cervical **Cervical Cancer**, Uterine Cervical Cancers, Uterine Uterine Cervical Cancers Cancer of Cervix Cervix Cancer Cancer, Cervix Cancers, Cervix) OR “Cervix Cancer*”[All Fields] OR “cervical cancer*”[All Fields] Or “cervix neoplasm*”[All Fields]

		OR “cancer* of the uterine cervix” [All Fields] Or “Cancer* or the cervix” [All Fields] OR “cervical neoplasm” [All Fields] OR “Cancer* of the Uterine Cervix” [All Fields] Or “Cancer* of Cervix” [All Fields]
Limits: from January 2009; humans		

3 EMBASE search strategy

“positron emission tomography”/syn OR “fluorodeoxyglucose f 18”/exp OR “fluorodeoxyglucose f 18”/syn OR “computer assisted emission tomography”/exp OR “computer assisted emission tomography” OR pet OR “pet scans” OR “pet scanner” OR “pet scan” OR “pet/ct scan” OR “pet/ct scans” OR “pet/ct” OR OR“positron emission tomography/computed tomography” OR OR pet NEAR/4 scan* OR pet NEAR/4 ct	AND	“cervical cancer”/syn OR “cervical cancers” OR “cervical neoplasm” OR “cervical neoplasms” OR “cervix cancer” OR “cervix cancers” OR “cervix neoplasm” OR “Cancer of the Uterine Cervix” OR “Cancer of the Cervix”

		OR Cervix NEAR/4 cancer* OR Cervical NEAR/4 neoplasm OR Cervical NEAR/4 cancer* OR Cervix NEAR/4 neoplasm
Limis: from January 2009; humans “article” OR “review”/it OR “short survey”		

APPENDIX 15

FDG-PET/CT for staging of testicular cancer

Characteristics of included studies

1 HTA report - KCE 2009 testicular cancer

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients;

	• Absence of adequate reference standard; • Absence of patient-based analysis; • Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	In the KCE report both the AHRQ 2008 report (AHRQ 2008) and Bourguet et al. (Bourguet 2007) identified one prospective study evaluating the use of PET for the staging of 46 patients having undergone orchidectomy and negative postoperative conventional staging. (Lassen 2003); histology/biopsy or clinical follow-up was used as reference standard. Sensitivity and specificity were 100% and 70% respectively for the detection of metastatic disease. In KCE 2009 one additional prospective study comparing PET and CT for the nodal staging of 72 patients with early-stage non-seminomatous germ cell tumours undergoing primary retroperitoneal lymph node dissection was identified; histopathology was used as reference standard. PET was found to be more sensitive (66% vs. 41%) and specific (97% vs. 95%) than CT, although the 95% confidence intervals were overlapping (de Wit 2008). The new evidence on the use of PET vs CT for the staging of testicular cancer consists of 1 primary study and is in line with the previous report. Overall, this evidence remains inconclusive.
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

PS - N staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Sterbis 2010	FDG-PET/CT	49	patients with histologically proven (post-orchiectomy) testicular cancer	93.3 (95% CI 66-99)	97.0 (95% CI 83-99)
	CT			60.0 (95% CI 33-82)	82.3 (95% CI 65-93)

PS - Sterbis 2010

Clinical features and settings	seminomatous germ cell tumors (SGCT) and non-seminomatous germ cell tumors (NSGCT) of the testis; country: USA
Participants	49 patients with histologically proven (post-orchiectomy) testicular cancer during the period 2003-2009 were included; median age (range): 27 (19-57); In total, retroperitoneal lymph node involvement was assessed in 28 patients. **Orchiectomy Pathology** No. Pts SGCT 15 NSGCT 34 **Clinical Stage** No. Pts Ia 22 Ib 1 Is 16 IIa 2 IIb 1 IIc 2 IIIa 2 IIIb 3 IIIc 0
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	retroperitoneal lymph node metastases. Reference standard*: pathological proof or clinical follow-up (or serum markers or CT size criteria).
Index and comparator tests	18 FDG-PET/CT; Comparators: CT, MRI.
Follow-up	Median follow-up: 39 months (range 4 to 85)

Notes	*For the purposes of calculating sensitivity, specificity a true positive was confirmed by histology obtained at retroperitoneal lymph node dissection (RPLND) (n = 3) or either positive serum markers or positive CT size criteria (n = 11) in those patients that did not undergo RPLND. A true negative was defined by pathology when available (n = 15) or by negative follow-up accompanying a negative PET/CT (n = 18). False positives (n=1) and negatives (n=1) were defined either by pathologic findings or clinical follow up contrary to initial PET/CT results.

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of 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** Could the selection of patients have introduced bias? **RISK:** LOW=YES/HIGH=NO/**UNCLEAR**
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? **Yes** If a threshold was used, was it pre-specified? **Unclear** Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? **Yes** Were the reference standard results interpreted without knowledge of the results of the index test? **Unclear** Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
4. FLOW AND TIMING (risk of bias)	High risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard:

		Was there an appropriate interval between index test(s) and reference standard? **Unclear** Did all patients receive a reference standard? **Yes** Did patients receive the same reference standard? **No** Were all patients included in the analysis? **Yes** Could the patient flow have introduced bias? RISK: LOW=YES/**HIGH=NO**/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? **CONCERN: LOW=YES**/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Unclear risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**

Characteristics of excluded studies

Boujelbene 2011

Reason for exclusion	Narrative review

Heidenreich 2010

Reason for exclusion	Clinical recommendations

Rioja 2010

Reason for exclusion	Narrative review

Sohaib 2011

Reason for exclusion	Narrative review

Zouhair 2010

Reason for exclusion	Narrative review

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”: ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. Testicular Neoplasms [Mesh explodes all trees]
10. Testicular Neoplasm*: ti,ab,kw
11. Testicular Cancer* : ti,ab,kw
12. Testis Cancer*: ti,ab,kw
13. Testis Neoplasm*:ti,ab,kw
14. Testicular NEAR/4 cancer*: ti,ab,kw
15. Testicular NEAR/4 neoplasm*: ti,ab,kw
16. Testis NEAR/4 cancer*: ti,ab,kw
17. Testis NEAR/4 cancer*: ti,ab,kw
18. Seminoma
19. Teratoma
20. 9/19 OR
21. 8 AND 20

Publication date: January 2009 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18” [Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg-pet”[All Fields]
17. “Positron-Emission Tomography” [Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. Testicular Neoplasms [Mesh explodes all trees]
26. “Testicular Neoplasm” [Title/Abstract]
27. “Testicular Neoplasms” [Title/Abstract]
28. “Testis Neoplasm” [Title/Abstract]
29. “Testis Neoplasms” [Title/Abstract]

30. “Testicular Cancer” [Title/Abstract]
31. “Testicular Cancers” [Title/Abstract]
32. “Testis Cancer” [Title/Abstract]
33. “Testis Cancers” [Title/Abstract]
34. Seminoma [Title/Abstract]
35. Teratoma [Title/Abstract]
36. Coriocarcinoma [Title/Abstract
37. 25/36 OR
38. 24 AND 37

Limit: Humans

Publication date: January 2009 - March 2012

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “positron emission tomography”/exp
3. “fluorodeoxyglucose f 18”/exp
4. “fluorodeoxyglucose f 18”/syn
5. “computer assisted emission tomography”/exp
6. “computer assisted emission tomography”/tw
7. pet/tw
8. “pet scans”/tw
9. “pet scanner”/tw
10. “pet scan”/tw
11. “pet/ct scan”/tw
12. “pet/ct scans”/tw
13. “pet/ct”/tw
14. “positron emission tomography/computed tomography”/tw
15. pet NEAR/4 scan*
16. pet NEAR/4 ct
17. 1/16 OR
18. “testicular Neoplasm”/de, syn, Keyword
19. “Testicular Neoplasms”/exp
20. “testicular cancer”/de, syn, Keyword”
21. “testicular cancers”/de, syn, Keyword”
22. “testis Neoplasms”/de, syn, Keyword
23. “testis cancer”/de, syn, Keyword
24. “Testis cancers”/de,syn;keyword
25. “testis neoplasm”/de,syn, keyword
26. “Testis neoplasms”/de,syn,Keyword
27. “testicular cancer”: ti, ab.
28. “testicular neoplasm”: ab:ti
29. “testicular neoplasms” : ab:ti
30. “testicular cancers”: :ab:ti
31. “testis cancers”: :ab:ti
32. “testis cancer”:ab:ti
33. “testis neoplasm”:ab:ti
34. “testis neoplasms”:ab:ti
35. Seminoma
36. Teratoma
37. Coriocarcinoma
38. 18/37 OR
39. 17 AND 38

Limit: Humans

Limit: “article” OR “review”/it OR “short survey”

Publication date: January 2009 - March 2012

APPENDIX 16

FDG-PET/CT for staging of prostate cancer

Characteristics of excluded studies

Budiharto 2011

Reason for exclusion	Contrast medium is choline

Contractor 2011

Reason for exclusion	Contrast medium is choline

De Visschere 2010

Reason for exclusion	Descriptive review

Garcia 2011

Reason for exclusion	Study assessing recurrence only

McCarthy 2011

Reason for exclusion	Not PET/CT

Panebianco 2012

Reason for exclusion	Contrast medium is choline

Souvatzoglu 2011

Reason for exclusion	Contrast medium is choline

Steuber 2010

Reason for exclusion	Contrast medium is choline

Watanabe 2010

Reason for exclusion	Study assessing diagnosis only, not tumour staging

Withofs 2011

Reason for exclusion	Contrast medium is (1)F-fluoride

Wurschmidt 2011

Reason for exclusion	Contrast medium is choline

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

“Positron-Emission Tomography” [MeSH descriptor explode all trees] OR “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees] OR. “positron emission tomography”: ti,ab,kw OR. pet*: ti,ab,kw OR. pet scan*: ti,ab,kw OR. “Fluorodeoxyglucose F18”: ti,ab,kw OR. fdg NEAR/2 18: ti,ab,kw	AND	Prostatic Neoplasms ”[Mesh explodes all trees] Prostate Neoplasms Neoplasms, Prostate Neoplasm, Prostate Prostate Neoplasm Neoplasms, Prostatic Neoplasm, Prostatic Prostatic Neoplasm **Prostate Cancer** Cancer, Prostate Cancers, Prostate Prostate Cancers Cancer of the Prostate Prostatic Cancer Cancer, Prostatic Cancers, Prostatic Prostatic Cancers Cancer of Prostate OR Prostat* Neoplasm*: ti,ab,kw OR Prostat* Cancer* : ti,ab,kw OR “Cancer* of Prostate” : ti,ab,kw OR Prostat* NEAR/4 cancer*: ti,ab,kw OR Prostat* NEAR/4 neoplasm*: ti,ab,kw

2 MEDLINE search strategy

"Fluorodeoxyglucose F18"[Mesh] **OR** "2-Fluoro-2-deoxyglucose" [All Fields] **OR** "18F Fluorodeoxyglucose" [All Fields] **OR** "F 18 Fluorodeoxyglucose" [All Fields] **OR** Fludeoxyglucose* [All Fields] **OR** "2 fluoro 2 deoxy d glucose"[All Fields] **OR** 18fluorodesoxyglucose*[All Fields] **OR** fluorodeoxyglucose*[All Fields] **OR** "fluorine 18 fluorodeoxyglucose" [All Fields] **OR** 18f dg*[All Fields]) **OR** 18fluorodeoxyglucose*[All Fields] **OR** 18fdg [All Fields] **OR** 18 fdg* [All Fields] **OR** fdg 18* [All Fields] **OR** fdg/* [All Fields] **OR** "fdg pet"[All Fields] **OR** "Positron-Emission Tomography"[Mesh] **OR** "positron emission tomography" [title/abstract] **OR** pet [title/abstract] **OR** "pet scan" [All Fields] **OR** "pet scans" [All Fields] **OR** "pet scanner" [All Fields] **OR** petscan [All Fields]	AND	Prostatic Neoplasms [Mesh explodes all trees] Prostate Neoplasms Neoplasms, Prostate Neoplasm, Prostate Prostate Neoplasm Neoplasms, Prostatic Neoplasm, Prostatic Prostatic Neoplasm **Prostate Cancer** Cancer, Prostate Cancers, Prostate Prostate Cancers Cancer of the Prostate Prostatic Cancer Cancer, Prostatic Cancers, Prostatic Prostatic Cancers Cancer of Prostate OR "Prostatic Neoplasm" [Title/Abstract] OR "Prostatic Cancer " [Title/Abstract] OR "Cancer of Prostate" [Title/Abstract] OR "Prostatic Neoplasms" [Title/Abstract] OR "Prostatic Cancers" [Title/Abstract]
Limits: from January 2009; humans		

3 EMBASE search strategy

"positron emission tomography"/syn OR "fluorodeoxyglucose f 18"/exp OR "fluorodeoxyglucose f 18"/syn OR "computer assisted emission tomography"/exp OR "computer assisted emission tomography" OR pet OR "pet scans" OR "pet scanner" OR "pet scan" OR "pet/ct scan" OR "pet/ct scans" OR "pet/ct" OR OR"positron emission tomography/computed tomography" OR OR pet NEAR/4 scan* OR pet NEAR/4 ct	AND	"Prostatic Neoplasms"/syn OR "Prostatic Neoplasms"/exp OR "prostatic cancer" OR "prostatic neoplasm" OR "prostate neoplasms" OR "prostate cancer" OR "Cancer of the Prostate" OR Prostatic NEAR/4 cancer* OR Prostatic NEAR/4 neoplasm
Limits: from January 2009; humans "article" OR "review"/it OR "short survey"		

APPENDIX 17

FDG-PET/CT for staging of penile cancer

Characteristics of included studies

HTA report - KCE 2009 penile cancer

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients; • Absence of adequate reference standard;

	• Absence of patient-based analysis; • Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	no systematic reviews or primary studies were found regarding penile cancer
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

SR - Sadeghi 2012 - N staging

Disease	penile squamous cell carcinoma
Index test	FDG-PET/CT
Comparators	not reported
Reference standard	inguinal lymph node dissection (or sentinel node biopsy) and/or follow-up
Target	inguinal regions lymph node metastases (N staging)
Studies included	7 diagnostic study with prospective or retrospective patients recruitment to perform an accuracy diagnostic test metanalysis.
Years covered by the search	up to May 2011 (submission date)
Comprehensive bibliographic search: at least two databases searched	Medline, SCOPUS, Google Scholar, Springer, and Science Direct
Characteristics of included studies clearly reported in tables	YES
Methodological quality of primary studies assessed; criteria reported	YES The Oxford Center for Evidence-Based Medicine checklist for diagnostic studies (http://www.cebm.net/index.aspx?o1025)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	YES
N. of included studies	7 primary diagnostic studies; One study had 2 separate subgroups of patients (cN and cN0 patients) that were included in the meta-analysis separately.
Design of included studies	diagnostic accuracy studies (not clearly specified for each study) with prospective or retrospective patients recruitment.
N. of included patients	115 patients (213 groins)
Diagnostic accuracy results (with heterogeneity)	**All patients** sensitivity : 80.9% (95% CI: 69.5%–89.4%) I^2:74.5% Cochrane Q/P: 27.41/<0.0003 specificity : 92.4% (95% CI: 86.8%–96.2%) I^2:61.2% Cochrane Q/P: 15.48/0.017 LR- : 0.288 (95% CI:0.094–0.878) I^2:86.4% Cochrane Q/P: 54.22/<0.0001

	LR+ : 6.461 (95% CI:2.088–19.993) I^2:59.7% Cochrane Q/P:14.90/0.021 DOR: 27.619 (95% CI:5.295–144.07) I^2: 57% Cochrane Q/P:13.94/0.03 AUC: 0.9089 **cN+ patients** sensitivity : 96.4% (95% CI: 81.7%–99.9%) specificity : 100% (95% CI: 83.9%–100%) LR- : 0.101 (95% CI:0.027–0.378) LR+ : 16.960 (95% CI:2.54–113.242) DOR: 229.20 (95% CI:17.743–2960.9) **cN0 patients** sensitivity : 56.5% (95% CI: 34.5%–76.8%) specificity : 85.9% (95% CI: 75.6%– 93.0%) LR- : 0.615 (95% CI:0.279–1.356) LR+ : 3.029 (95% CI:1.510–6.078) DOR: 7.532 (95% CI:2.040–27.808)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? **Unclear** Was a case-control design avoided? **Unclear** Did the study avoid inappropriate exclusions? **Unclear** Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? **Yes** If a threshold was used, was it pre-specified? **Unclear** Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**

3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? **Unclear** Were the reference standard results interpreted without knowledge of the results of the index test? **Unclear** Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? **Yes** Did all patients receive a reference standard? **Yes** Did patients receive the same reference standard? **UNCLEAR** Were all patients included in the analysis? **Yes** Could the patient flow have introduced bias? RISK: **LOW=YES**/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN**: LOW=YES**/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR

PS - N staging

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Souillac 2012	FDG-PET/CT	30	squamous cell carcinoma (SCC) of the penis all N stages	91	89.8
		22	squamous cell carcinoma (SCC) of the penis cN0	75	87.5
		8	squamous cell carcinoma (SCC) of the penis cN+	100	100
	CT	30	squamous cell carcinoma (SCC) of the penis all N stages	91	81.6
		22	squamous cell carcinoma (SCC) of the penis cN0	100	77.5
		8	squamous cell carcinoma (SCC) of the penis cN+	85.7	100

PS - Souillac 2012

Clinical features and settings	squamous cell carcinoma (SCC) of the penis; country: France
Participants	30 patients with histologically proven penile carcinoma during the period from March 2005 until January 2010 were included; mean age (range): 69 (41-94); In total, lymph node involvement was assessed in 60 inguinal groins. pT* stage 1b (n):12 pT stage 2 (n): 13 pT stage 3 (n): 5 cN stage 0 (n): 22 cN stage 1 (n): 6 cN stage 2 (n): 2 pN stage 0 (n): 21 pN stage 1 (n): 7 pN stage 2 (n): 2 * T1b: Tumour invades subepithelial connective tissue without with lymphovascular invasion or is poorly differentiated or undifferentiated (T1G3-4) according to UICC TNM 2009 classification. pT: pathologic classification of the primary tumor status

	cN: clinically classification of the lymph node inguinal region pN: pathologic classification of the lymph node inguinal region
Study design	diagnostic cross sectional study with prospective recruitment
Target condition and reference standard(s)	inguinal regions lymph node metastases; Reference standard: pathological proof.
Index and comparator tests	FDG-PET/CT; comparator: clinical examination, CT.
Follow-up	not stated
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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** Could the selection of patients have introduced bias? **RISK: LOW=YES**/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? **Yes** If a threshold was used, was it pre-specified? **Unclear** Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? **Unclear** Were the reference standard results interpreted without knowledge of the results of the index test? **Unclear** Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/**UNCLEAR**

4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? **Yes** Did all patients receive a reference standard? **Yes** Did patients receive the same reference standard? **Unclear** Were all patients included in the analysis? **Yes** Could the patient flow have introduced bias? RISK: **LOW=YES**/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: **LOW=YES**/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? **CONCERN: LOW=YES**/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Unclear risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/**UNCLEAR**

Characteristics of excluded studies

Graafland 2009

Reason for exclusion	Retrospective study design

Hughes 2009

Reason for exclusion	Narrrative review

Johnson 2009

Reason for exclusion	Case report

Rosevear 2011

Reason for exclusion	Retrospective study design; only 3 patients included.

Scher 2005

Reason for exclusion	Only 8 patients included; It was not considered in KCE 2009.

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. “positron emission tomography”: ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. “Fluorodeoxyglucose F18”: ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. Penile Neoplasms [Mesh explodes all trees]
10. Penile Neoplasm*: ti,ab,kw
11. Penile Cancer* : ti,ab,kw
12. Penile NEAR/4 cancer*: ti,ab,kw
13. Penile NEAR/4 neoplasm*: ti,ab,kw
14. 9/13 OR
15. 8 AND 14

Publication date: January 2009 - March 2012

2 MEDLINE search strategy

1. “Fluorodeoxyglucose F18” [Mesh]
2. “2-Fluoro-2-deoxyglucose” [All Fields]
3. “18F Fluorodeoxyglucose” [All Fields]
4. “F 18 Fluorodeoxyglucose” [All Fields]
5. Fludeoxyglucose* [All Fields]
6. “2 fluoro 2 deoxy d glucose”[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. “fluorine 18 fluorodeoxyglucose” [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. “fdg-pet”[All Fields]
17. “Positron-Emission Tomography” [Mesh]
18. “positron emission tomography” [title/abstract]
19. pet [title/abstract]
20. “pet scan” [All Fields]
21. “pet scans” [All Fields]
22. “pet scanner” [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. Penile Neoplasms [Mesh explodes all trees]
26. “Penile Neoplasm” [Title/Abstract]
27. “Penile Neoplasms” [Title/Abstract]
28. “Penile Cancer” [Title/Abstract]
29. “Penile Cancers” [Title/Abstract]
30. 25/29 OR
31. 24 AND 30

Limit: Humans

Publication date: January 2009 - March 2012

3 EMBASE search strategy

1. “positron emission tomography”/syn
2. “positron emission tomography”/exp
3. “fluorodeoxyglucose f 18”/exp
4. “fluorodeoxyglucose f 18”/syn
5. “computer assisted emission tomography”/exp
6. “computer assisted emission tomography”/tw
7. pet/tw
8. “pet scans”/tw
9. “pet scanner”/tw
10. “pet scan”/tw
11. “pet/ct scan”/tw
12. “pet/ct scans”/tw
13. “pet/ct”/tw
14. “positron emission tomography/computed tomography”/tw
15. pet NEAR/4 scan*
16. pet NEAR/4 ct
17. 1/15 OR
18. “Penile Neoplasm”/de, syn, Keyword
19. “Penile Neoplasms”/exp
20. “penile cancer”/de, syn, Keyword”
21. “penile cancers”/de, syn, Keyword”
22. “PenileNeoplasms”/de, syn, Keyword
23. “penile cancer”: ti, ab.
24. “penile neoplasm”: ab:ti
25. “penile adenocarcinoma”: ab:ti
26. “penile neoplasms”: :ab:ti
27. “penile cancers”: :ab:ti
28. Penile NEAR/4 cancer*
29. Penile NEAR/4 neoplasm
30. Penile NEAR/4 cancers
31. Penile NEAR/4 neoplasms
32. 18/31 OR
33. 17 AND 32

Limit: Humans

Limit: “article” OR “review”/it OR “short survey”

Publication date: January 2009 - March 2012

APPENDIX 18

FDG-PET/CT for staging of melanoma

Characteristics of included studies

HTA report - KCE 2009

Document ID	KCE 2009
Objectives	To answer the following research questions: What is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT?
Methods	Fifteen HTA agency databases were searched in addition to an OVID Medline search limited to articles published in English, French or Dutch published between 2005 and 2009. The criteria for inclusion were: systematic reviews and prospective and retrospective primary studies of diagnostic performance of PET or PET/CT in people with malignancies. Retrospective studies design or the presence of differential verification (i.e. more than one reference standard used) were not exclusion criteria as such. Incorporation bias (i.e. the use of the index test as a part of the reference standard) was not used as an exclusion criterion, but was considered a criterion of low quality. Prognostic studies without a multivariate analysis and using the index test to modify the management were excluded. Editorials, letters and case reports were excluded. There was a quality threshold to inclusion as reports, systematic reviews and meta-analyses not reporting the search strategy or quality assessment were excluded. “For diagnostic accuracy studies we used the following exclusion criteria: • Inability to reconstruct the contingency table(s); • Sample size (i.e. total number of subjects) < 20 patients; • Absence of adequate reference standard;

	• Absence of patient-based analysis; • Case-control study design; • Presence of partial verification (i.e. part of the population not receiving verification with the reference standard)". Quality was assessed as follows: for HTA reports the INAHTA checklist, for systematic reviews, prognostic studies and RCTs the relevant Dutch Cochrane Centre checklist for diagnostic studies the QUADAS checklist The tests were assessed by tumour by their technical accuracy, place in clinical pathway, diagnostic accuracy, impact on patient outcome and cost-effectiveness. On this basis the performance in each tumor was graded as: • Level 1: Technical accuracy • Level 2: Diagnostic accuracy • Level 3: Impact on patient outcome • Level 4: Cost-effectiveness
Conclusions	Evidence consistently shows a low sensitivity for the detection of lymph node metastasis in cN0 melanomas (level 2). It also found that a good balance between sensitivity and specificity in advanced stages for the detection of distant metastasis in patients with primary and recurrent malignant melanoma (level 2).
Notes	This review has several problems. The inclusion criteria are unclear and have been precied partly by guesswork. Methodological quality assessment relied in two cases instrumentswhich are not designed for quality assessment (the INAHTA checklist) or do not address general study design points (QUADAS).

SR - Xing 2011 – N and M staging

Disease	melanoma
Index test	FDG-PET/CT
Comparators	PET, CT, ultrasonography
Reference standard	N staging: sentinel lymph node biopsy with pathological confirmation M staging: histological analysis and/or 6 months follow-up
Target	regional lymph node staging distanta metastasis staging
Studies included	mixed retrospective and prospective design studies
Years covered by the search	up to June 2009
Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE, EMBASE, Cancerlit and the Controlled Trials Register from the Cochrane Library)
Characteristics of included studies clearly reported in tables	Yes
Methodological quality of primary studies assessed; criteria reported	Yes (QUADAS was used to assess quality of included studies; quality scores reported as scattergrams)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes (Bayesian binomial markov modelling. Likelihood of publication bias was assessed by funnel plots visual inspection)
N. of included studies	74 (21 ultrasonography, 13 CT, 45 FDG-PET, 13 FDG-PET/CT). Patients were enrolled exclusively for the purposes of primary staging in 30 studies or surveillance in 34 studies
Design of included studies	most studies included in this meta-analysis had a retrospective design
N. of included patients	mean number of participants per study 140 (range = 10–2008 participants per study)
Diagnostic accuracy results (with heterogeneity)	Regional lymph nodes staging (median) ultrasonography sensitivity 60% (95% CI 33-83%)

	specificity 97% (95% CI 88-99%) CT sensitivity 9% (95% CI 1-52%) specificity 92% (95% CI 50-99%) FDG-PET sensitivity 30% (95% CI 12-55%) specificity 96% (95% CI 87-99%) FDG-PET/CT sensitivity 11% (95% CI 1-50%) specificity 97% (95% CI 78-100%) Distant metastases (median) CT sensitivity 51% (95% CI 24-76%) specificity 69% (95% CI 30-92%) FDG-PET sensitivity 74% (95% CI 51-88%) specificity 75% (95% CI 45-91%) FDG-PET/CT sensitivity 80% (95% CI 53-93%) specificity 87% (95% CI 54-97%)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	High risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? majority of studies with retrospective design Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the

		results of the reference standard? Uncler If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Unclear Did patients receive the same reference standard? Unclear Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Unclear risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

Characteristics of excluded studies

Bastiaannet 2009

Reason for exclusion	PET and CT were not performed together but in random order at two different sites (with a median of 5 days between the two). The results were compared as a head to head

Bastiaannet 2011

Reason for exclusion	Economic analysis bolted onto Bastiaannet 2009. Exclude for the same reasons

Camargo Etchebehere 2010

Reason for exclusion	Non comparative

Dellestable 2011

Reason for exclusion	No reference standard

Heusner 2011

Reason for exclusion	Mixed tumors with no data breakdown reported

Jimenez-Requena 2010

Reason for exclusion	Search period earlier then KCE review

Peric 2011

Reason for exclusion	No reference test

Ribas 2011

Reason for exclusion	No reference standard. The study reports a comparison of different probes

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

“Positron-Emission Tomography” [MeSH descriptor explode all trees] OR “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees] OR. “positron emission tomography”: ti,ab,kw OR. pet*: ti,ab,kw OR. pet scan*: ti,ab,kw OR. “Fluorodeoxyglucose F18”: ti,ab,kw OR. fdg NEAR/2 18: ti,ab,kw	AND	Melanoma ”[Mesh explodes all trees] contiene: Melanomas Malignant **Melanoma** Malignant Melanomas **Melanoma**, Malignant Melanomas, Malignant OR Melanom*: ti,ab,kw

Publication date: January 2009 - March 2012

2 MEDLINE search strategy

“Fluorodeoxyglucose F18”[Mesh] **OR** “2-Fluoro-2-deoxyglucose” [All Fields] **OR** “18F Fluorodeoxyglucose” [All Fields] **OR** “F 18 Fluorodeoxyglucose” [All Fields] **OR** Fludeoxyglucose* [All Fields] **OR** “2 fluoro 2 deoxy d glucose”[All Fields] **OR** 18fluorodesoxyglucose*[All Fields] **OR** fluorodeoxyglucose*[All Fields] **OR** “fluorine 18 fluorodeoxyglucose” [All Fields] **OR** 18f dg*[All Fields]) **OR** 18fluorodeoxyglucose*[All Fields] **OR** 18fdg [All Fields] **OR** 18 fdg* [All Fields] **OR** fdg 18* [All Fields] **OR** fdg/* [All Fields] **OR** “fdg pet”[All Fields] **OR** “Positron-Emission Tomography”[Mesh] **OR** “positron emission tomography” [title/abstract] **OR** pet [title/abstract] **OR**	AND	Melanoma ”[Mesh explodes all trees] includes: Melanomas Malignant **Melanoma** Malignant Melanomas **Melanoma**, Malignant Melanomas, Malignant OR Melanom*:[All Fields]

“pet scan” [All Fields] **OR** “pet scans” [All Fields] **OR** “pet scanner” [All Fields] **OR** petscan [All Fields]		

Limit: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2009 - March 2012

3 EMBASE search strategy

“Fluorodeoxyglucose F18”[Mesh] **OR** “2-Fluoro-2-deoxyglucose” [All Fields] **OR** “18F Fluorodeoxyglucose” [All Fields] **OR** “F 18 Fluorodeoxyglucose” [All Fields] **OR** Fludeoxyglucose* [All Fields] **OR** “2 fluoro 2 deoxy d glucose”[All Fields] **OR** 18fluorodesoxyglucose*[All Fields] **OR** fluorodeoxyglucose*[All Fields] **OR** “fluorine 18 fluorodeoxyglucose” [All Fields] **OR** 18f dg*[All Fields]) **OR** 18fluorodeoxyglucose*[All Fields] **OR** 18fdg [All Fields] **OR** 18 fdg* [All Fields] **OR** fdg 18* [All Fields] **OR** fdg/* [All Fields] **OR** “fdg pet”[All Fields] **OR** “Positron-Emission Tomography”[Mesh] **OR** “positron emission tomography” [title/abstract] **OR** pet [title/abstract] **OR** “pet scan” [All Fields] **OR** “pet scans” [All Fields] **OR** “pet scanner” [All Fields] **OR** petscan [All Fields]	AND	Melanoma: ab,ti OR Melanomas: ab,ti OR Melanoma: de,syn, keyword OR Melanomas: de,syn, keyword

Limit: Humans; “article” OR “review”/it OR “short survey”
Languages: English, French, Italian, Spanish
Publication date: January 2009 - March 2012

APPENDICES to FDG-PET/CT for staging of osteosarcomas

Characteristics of excluded studies

Bastianneet 2004

Reason for exclusion	Systematic review with very poor methods (no details of meta-analysis and unclear inclusion criteria) and evaluating PET performance only

Portwine 2010

Reason for exclusion	Descriptive review

Volker

Reason for exclusion	PET only study

Ye 2008

Reason for exclusion	PET only study

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

“Positron-Emission Tomography” [MeSH descriptor explode all trees] OR “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees] OR. “positron emission tomography”: ti,ab,kw OR. pet*: ti,ab,kw OR. pet scan*: ti,ab,kw OR. “Fluorodeoxyglucose F18”: ti,ab,kw OR. fdg NEAR/2 18: ti,ab,kw	AND	Osteosarcoma ”[Mesh explodes all trees] contiene: Osteosarcomas Osteosarcoma Tumor Osteosarcoma Tumors Tumor, Osteosarcoma Tumors, Osteosarcoma Sarcoma, Osteogenic Osteogenic Sarcoma Osteogenic Sarcomas Sarcomas, Osteogenic OR Osteorsarcom*: ti,ab,kw OR Sarcoma* NEAR Osteogenic: ti,ab,kw OR Osteogenic NEAR Sarcoma* ti,ab,kw OR Bone NEAR Sarcoma*: ti,ab,kw

2 MEDLINE search strategy

“Fluorodeoxyglucose F18”[Mesh] **OR** “2-Fluoro-2-deoxyglucose” [All Fields] **OR** “18F Fluorodeoxyglucose” [All Fields] **OR** “F 18 Fluorodeoxyglucose” [All Fields] **OR** Fludeoxyglucose* [All Fields] **OR** “2 fluoro 2 deoxy d glucose”[All Fields] **OR** 18fluorodesoxyglucose*[All Fields] **OR** fluorodeoxyglucose*[All Fields] **OR** “fluorine 18 fluorodeoxyglucose” [All Fields] **OR** 18f dg*[All Fields]) **OR** 18fluorodeoxyglucose*[All Fields] **OR** 18fdg [All Fields] **OR** 18 fdg* [All Fields] **OR** fdg 18* [All Fields] **OR** fdg/* [All Fields] **OR** “fdg pet”[All Fields] **OR** “Positron-Emission Tomography”[Mesh] **OR** “positron emission tomography” [title/abstract] **OR**	AND	Osteosarcoma ”[Mesh explodes all trees] contiene: Osteosarcomas Osteosarcoma Tumor Osteosarcoma Tumors Tumor, Osteosarcoma Tumors, Osteosarcoma Sarcoma, Osteogenic Osteogenic Sarcoma Osteogenic Sarcomas Sarcomas, Osteogenic OR Osteorsarcom*: ti,ab,kw OR Sarcoma* NEAR Osteogenic: ti,ab,kw

pet [title/abstract] **OR** “pet scan” [All Fields] **OR** “pet scans” [All Fields] **OR** “pet scanner” [All Fields] **OR** petscan [All Fields]		OR Osteogenic NEAR Sarcoma* ti,ab,kw OR Bone NEAR Sarcoma*: ti,ab,kw
Limits: from January 2006; humans		

3 EMBASE search strategy

“Fluorodeoxyglucose F18”[Mesh] **OR** “2-Fluoro-2-deoxyglucose” [All Fields] **OR** “18F Fluorodeoxyglucose” [All Fields] **OR** “F 18 Fluorodeoxyglucose” [All Fields] **OR** Fludeoxyglucose* [All Fields] **OR** “2 fluoro 2 deoxy d glucose”[All Fields] **OR** 18fluorodesoxyglucose*[All Fields] **OR** fluorodeoxyglucose*[All Fields] **OR** “fluorine 18 fluorodeoxyglucose” [All Fields] **OR** 18f dg*[All Fields]) **OR** 18fluorodeoxyglucose*[All Fields] **OR** 18fdg [All Fields] **OR** 18 fdg* [All Fields] **OR** fdg 18* [All Fields] **OR** fdg/* [All Fields] **OR** “fdg pet”[All Fields] **OR** “Positron-Emission Tomography”[Mesh] **OR** “positron emission tomography” [title/abstract] **OR** pet [title/abstract] **OR** “pet scan” [All Fields] **OR** “pet scans” [All Fields] **OR** “pet scanner” [All Fields] **OR** petscan [All Fields]	AND	Osteorsarcoma: ab,ti OR Osteorsarcomas: ab,ti OR “Sarcoma Osteogenic”: ab,ti OR “Sarcomas Osteogenic”: ab,ti OR “Osteogenic Sarcoma” ab,ti OR “Osteogenic Sarcomas” ab,ti OR “Bone Sarcoma” ab,ti OR “Bone Sarcomas” ab,ti OR Osteorsarcoma: de,syn, keyword OR Osteorsarcomas: de,syn, keyword OR “Sarcoma Osteogenic”: de,syn, keyword OR “Sarcomas Osteogenic”: de,syn, keyword

		OR “Osteogenic Sarcoma” de,syn, keyword OR “Osteogenic Sarcomas” de,syn, keyword OR “Bone Sarcoma” de,syn, keyword OR “Bone Sarcomas” de, syn, keyword
Limits: from January 2006; humans “article” OR “review”/it OR “short survey”		

APPENDIX 19

FDG-PET/CT for staging of Hodgkin's lymphoma

Characteristics of included studies

HTA report - ASSR Lymphomas 2012

Document ID	ASSR-RER 2012 - Lymphomas
Objectives	to define criteria for appropriate use of FDG-PET for patients with Hodgkin's lymphoma
Methods	A panel of experts working in Health Trusts and Teaching Hospitals of Emilia-Romagna was convened to discuss and agree on the methodology for a research programme aimed at defining the criteria for appropriate use of PET in Hodgkin's lymphoma. On the basis of the clinical pathway of patients with Hodgkin's lymphoma the panel examined and assessed the role of FDG-PET for 6 clinical indications (staging, dose painting definition in involved-field radiation treatment, during treatment evaluation of early response to therapy, end of treatment evaluation of response to therapy, follow up of patients with no suspicion of recurrence, staging of recurrence). The following databases were searched for the period between January 2006 and February 2011: Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE - The Cochrane Library); Health Technology Assessment Database; Cochrane Central Register of Controlled Trials; National Library of Medicine's Medline database (PubMed); Elsevier's Embase. Language restrictions: English, Italian, French and Spanish. Selection criteria Type of studies: systematic reviews, RCTs, CCTs, cross-sectional diagnostic studies, prospective or retrospective cohort studies, case series of at least 10 patients Participants: patients with breast cancer Intervention: FDG-PET or CT/PET Reference standard: histology or clinical follow up Comparator: any other imaging technique Outcomes> sensitivity, specificity, LR, metabolic/tumor response, time to recurrence, local, local-regional and distant recurrence, disease free survival, disease survival, overall survival Assessment of methodological quality of studies The following criteria have been used for the quality assessment of different study designs. Systematic reviews: criteria drawn from the AMSTAR checklist Diagnostic cross sectional studies: criteria drawn from the QUADAS checklist Randomized controlled trials: criteria suggested by the Cochrane Handbook Case control studies and cohort studies: criteria drawn from the New Castle-Ottawa checklist Level of evidence for estimates of diagnostic accuracy were assigned according to GRADE

	categorization of the quality of evidence Each member of the panel voted the level of appropriateness for each clinical question. Two rounds of votes were requested for the judgment of appropriateness and results were analysed using the RAND/UCLA Appropriateness Method. The use of FDG-PET for a specific clinical indication was judged was judged as *appropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 7-9 score region as *inappropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 1-3 score region. Finally the use of FDG-PET was judged as *uncertain* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 4-6 score region or when no agreement was reached after the second round of voting. Clinical indications for which the panel does not reach an agreement on level of appropriateness after two rounds of voting also fall in the *uncertain* category.
Conclusions	STAGING OF HODGKIN'S LYMPHOMA - APPROPRIATE During the first meeting the panel reached an agreement in judging appropriate the use of FDG-PET for staging patients with Hodgkin's lymphoma, in order to distinguish early, localised stage (I and II) from advanced, extended (stage III and IV) disease and direct patients to most appropriate treatment. The level of evidence for estimates of FDG-PET diagnostic accuracy was moderate, with FDG-PET performing better than comparator for detection of both linfonodal and extra-nodal involvement.
Notes	Meta-analysis of diagnostic accuracy estimates was not performed

PS - Staging (bone marrow disease)

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Pelosi 2011	FDG-PET/CT	130	patients at initial staging with Hodgkin's lymphoma	78.6	100
	bone marrow biopsy			42.9	100
Purz 2011	FDG-PET/CT	175	pediatric patients at initial staging with Hodgkin's lymphoma	100	77.3

PS - Pelosi 2011

Clinical features and settings	Hodgkin's lymphoma, aggressive non Hodgkin's; Country: Italy
Participants	337 consecutive patients with Hodgkin's lymphoma (130) or aggressive non Hodgkin's patients (207; diffuse large B-cell 120, follicular grade II-III 48, mantle cell 7, others 32). Age (yr) median 49.4, range 11–84; male 189, female 148
Study design	prospective diagnostic accuracy study, with consecutive accrual
Target condition and reference standard(s)	Bone marrow extension of disease Reference standard: composite: bone marrow biopsy and imaging follow up (with FDG-PET/CT and /or MRI)
Index and comparator tests	FDG-PET/CT comparator: bone marrow biopsy
Follow-up	not reported
Notes	data reported also for disease subgroup

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

2A. INDEX TEST(S) (risk of bias)	High risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? No (incorporation bias) If a threshold was used, was it pre-specified? Yes Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	High risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? No (incorporation bias) Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

3B. REFERENCE STANDARD (concern of applicability)	High risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Purz 2011

Clinical features and settings	Hodgkin's lymphoma; Country: Germany
Participants	175 pediatric patients with newly diagnosed classical Hodgkin's lymphoma (130) with stage greater than IIA.
Study design	prospective diagnostic accuracy study; not consecutive accrual
Target condition and reference standard(s)	Bone marrow extension of disease Reference standard: composite: bone marrow biopsy
Index and comparator tests	FDG-PET/CT or FDG-PET
Follow-up	not reported
Notes	

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	High risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? No (non consecutive enrollment) Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Unclear risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Yes Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR

3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index test? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Unclear Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	High risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

Characteristics of excluded studies

Abdulqadhr 2011

Reason for exclusion	not available diagnostic accuracy estimates (sensitivity, specificity)

Cerci 2011

Reason for exclusion	assessment of FDG-PET

Cheng 2011

Reason for exclusion	retrospective study

Gu 2011

Reason for exclusion	FDG-PET as reference standard, lesion-based analysis

Huang 2011

Reason for exclusion	retrospective study

Ilica 2011

Reason for exclusion	narrative review

Mittal 2011

Reason for exclusion	unclear design of study

Paulino 2011

Reason for exclusion	unclear design of study, nodal based analysis

van Ufford 2011

Reason for exclusion	agreement study

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. "positron emission tomography":ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. "Fluorodeoxyglucose F18": ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. “Lymphoma”/exp
10. 8 AND 9

Publication date: January 2011 - December 2011

2 MEDLINE search strategy

1. "Fluorodeoxyglucose F18"[Mesh]
2. "2-Fluoro-2-deoxyglucose" [All Fields]
3. "18F Fluorodeoxyglucose" [All Fields]
4. "F 18 Fluorodeoxyglucose" [All Fields]
5. Fludeoxyglucose* [All Fields]
6. "2 fluoro 2 deoxy d glucose"[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. "fluorine 18 fluorodeoxyglucose" [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. "fdg pet"[All Fields]
17. "Positron-Emission Tomography"[Mesh]
18. "positron emission tomography" [title/abstract]
19. pet [title/abstract]
20. "pet scan" [All Fields]
21. "pet scans" [All Fields]
22. "pet scanner" [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. "Lymphoma"[Mesh:noexp]
26. "Hodgkin Disease"[Mesh]
27. "Lymphoma, Non-Hodgkin"[Mesh]
28. "lymphomas"[Title/Abstract]
29. "lymphoma"[Title/Abstract]
30. "hodgkin's"[Title/Abstract]
31. "hodgkins"[Title/Abstract]
32. "hodgkin"[Title/Abstract]
33. "hodgkin s"[Title/Abstract]
34. "lymphogranuloma"[Title/Abstract]
35. "non hodgkin"[Title/Abstract]
36. "non hodgkin s b"[Title/Abstract]
37. "non hodgkin's"[Title/Abstract]
38. "reticulum cell sarcoma"[Title/Abstract]
39. "reticulum cell sarcomas"[Title/Abstract]
40. "reticulosarcoma"[Title/Abstract]
41. "reticulosarcomas"[Title/Abstract]
42. "lymphosarcoma"[Title/Abstract]
43. "lymphosarcomas"[Title/Abstract]

44. "lymphatic sarcoma"[Title/Abstract]
45. "lymphatic sarcomas"[Title/Abstract]
46. "burkitt's"[Title/Abstract]
47. "burkitt"[Title/Abstract]
48. "burkitt s"[Title/Abstract]
49. "lymphocytic leukemia"[Title/Abstract]
50. "lymphocytic leukemias"[Title/Abstract]
51. "lymphomatoid granulomatoses"[Title/Abstract]
52. "lymphomatoid granulomatosis"[Title/Abstract]
53. "brill symmers disease"[Title/Abstract]
54. "immunoblastoma"[Title/Abstract]
55. "immunoblastomas"[Title/Abstract]
56. "immunoblastosarcoma"[Title/Abstract]
57. "immunoblastosarcomas"[Title/Abstract]
58. "immunoblastic sarcoma"[Title/Abstract]
59. "immunoblastic sarcomas"[Title/Abstract]
60. "granulomatous slack skin"[Title/Abstract]
61. "lymphomatoid papulosis"[Title/Abstract]
62. "mycosis fungoides"[Title/Abstract]
63. "pagetoid reticulosis"[Title/Abstract]
64. "woringer kolopp disease"[Title/Abstract]
65. "ketron goodman disease"[Title/Abstract]
66. "sezary lymphoma"[Title/Abstract]
67. "sezary's syndrome"[Title/Abstract]
68. "sezary syndrome"[Title/Abstract]
69. 25/68 OR
70. 69 AND 24
71. "editorial"[Publication Type]
72. "comment"[Publication Type]
73. "letter"[Publication Type]
74. 71/73 OR
75. 70 NOT 74

LIMITS: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2011 - December 2011

3 EMBASE search strategy

1. 'positron emission tomography'/syn
2. 'fluorodeoxyglucose f 18'/exp
3. 'fluorodeoxyglucose f 18'/syn
4. 'computer assisted emission tomography'/exp
5. 'computer assisted emission tomography' OR
6. pet
7. 'pet scans'
8. 'pet scanner'
9. 'pet scan'
10. 'pet/ct scan'
11. 'pet/ct scans'
12. 'pet/ct'
13. 'positron emission tomography/computed tomography'
14. pet NEAR/4 scan*
15. pet NEAR/4 ct
16. 1/23 OR
17. 'lymphoma'/de
18. 'hodgkin disease'/exp
19. 'classical hodgkin lymphoma'/exp
20. 'nonhodgkin lymphoma'/exp
21. 'intestine lymphoma'/de
22. 'skin lymphoma'/exp
23. 'reed sternberg cell'/de
24. 'cutaneous t cell lymphoma'/exp

25. 'histiocytic lymphoma'/exp
26. 'marginal zone lymphoma'/exp
27. 't cell lymphoma'/exp
28. 17/27 OR
29. 16 AND 28
LIMITS:
Publication Type: article; article in press; erratum; short survey
Humans
Languages: English, French, Italian, Spanish
Publication date: January 2011 - December 2011

APPENDIX 20

FDG-PET/CT for staging of aggressive non-Hodgkin's lymphoma

Characteristics of included studies

HTA report - ASSR Lymphomas 2012

Document ID	ASSR-RER 2012 - Lymphomas
Objectives	to define criteria for appropriate use of FDG-PET for patients with non-Hodgkin's lymphoma
Methods	A panel of experts working in Health Trusts and Teaching Hospitals of Emilia-Romagna was convened to discuss and agree on the methodology for a research programme aimed at defining the criteria for appropriate use of PET in non-Hodgkin's lymphoma. On the basis of the clinical pathway of patients with Hodgkin's lymphoma the panel examined and assessed the role of FDG-PET for 6 clinical indications (staging, dose painting definition in involved-field radiation treatment, during treatment evaluation of early response to therapy, end of treatment evaluation of response to therapy, follow up of patients with no suspicion of recurrence, staging of recurrence). The following databases were searched for the period between January 2006 and February 2011: Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE - The Cochrane Library); Health Technology Assessment Database; Cochrane Central Register of Controlled Trials; National Library of Medicine's Medline database (PubMed); Elsevier's Embase. Language restrictions: English, Italian, French and Spanish. Selection criteria Type of studies: systematic reviews, RCTs, CCTs, cross-sectional diagnostic studies, prospective or retrospective cohort studies, case series of at least 10 patients Participants: patients with breast cancer Intervention: FDG-PET or CT/PET Reference standard: histology or clinical follow up Comparator: any other imaging technique Outcomes> sensitivity, specificity, LR, metabolic/tumor response, time to recurrence, local, local-regional and distant recurrence, disease free survival, disease survival, overall survival Assessment of methodological quality of studies The following criteria have been used for the quality assessment of different study designs. Systematic reviews: criteria drawn from the AMSTAR checklist Diagnostic cross sectional studies: criteria drawn from the QUADAS checklist Randomized controlled trials: criteria suggested by the Cochrane Handbook Case control studies and cohort studies: criteria drawn from the New Castle-Ottawa checklist

	Level of evidence for estimates of diagnostic accuracy were assigned according to GRADE categorization of the quality of evidence Each member of the panel voted the level of appropriateness for each clinical question. Two rounds of votes were requested for the judgment of appropriateness and results were analysed using the RAND/UCLA Appropriateness Method. The use of FDG-PET for a specific clinical indication was judged was judged as *appropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 7-9 score region as *inappropriate* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 1-3 score region. Finally the use of FDG-PET was judged as *uncertain* when, after discarding one extreme high and one extreme low rating, all remaining ratings fell within the 4-6 score region or when no agreement was reached after the second round of voting. Clinical indications for which the panel does not reach an agreement on level of appropriateness after two rounds of voting also fall in the *uncertain* category.
Conclusions	STAGING OF NON-HODGKIN'S LYMPHOMA - APPROPRIATE During the first meeting the panel reached an agreement in judging appropriate (median score 8; range 7-9) the use of FDG-PET for staging patients with aggressive non-Hodgkin's lymphoma, in order to distinguish early, localised stage (I and II) from advanced, extended (stage III and IV) disease and direct patients to most appropriate treatment. The level of evidence for estimates of FDG-PET diagnostic accuracy was moderate, with FDG-PET performing better than comparators for detection of both linfonodal/extra-nodal involvement and bone marrow involvement.
Notes	Meta-analysis of diagnostic accuracy estimates was not performed

SR - Chen 2011

Disease	aggressive non-Hodgkin's lymphoma, indolent non-Hodgkin's lymphoma
Index test	FDG-PET, FDG-PET/CT
Comparators	none
Reference standard	histopathology, bone marrow histology
Target	diagnostic accuracy for staging (bone marrow infiltration)
Studies included	diagnostic accuracy studies with prospective or retrospective set-up of study
Years covered by the search	up to May 2010
Comprehensive bibliographic search: at least two databases searched	Yes (MEDLINE; EBM review)
Characteristics of included studies clearly reported in tables	No
Methodological quality of primary studies assessed; criteria reported	Yes (Cochrane diagnostic group tool)
Meta-analysis performed with appropriate statistic methods (including heterogeneity evaluation)	Yes
N. of included studies	aggressive non-Hodgkin's lymphoma 6 studies indolent non-Hodgkin's lymphoma 3 studies
Design of included studies	4 studies retrospective design; 4 studies prospective design
N. of included patients	aggressive non-Hodgkin's lymphoma FDG-PET 134 patients FDG-PET/CT 237 patients indolent non-Hodgkin's lymphoma FDG-PET or FDG-PET/CT 156 patients
Diagnostic accuracy results (with heterogeneity)	aggressive non-Hodgkin's lymphoma FDG-PET/CT Sensitivity 74.0% (95% CI 65.0-83.0%)

	Specificity 80.0% (95% CI 74.0-87.0%) FDG-PET Sensitivity 74.0% (95% CI 62.0-86.0%) Specificity 92.0% (95% CI 86.0-98.0%)

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Unclear risk	Describe methods of patient selection: Was a consecutive or random sample of patients enrolled? Unclear (50% of studies retrospective design and non consecutive enrollment) Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	Low risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? Yes If a threshold was used, was it pre-specified? Yes Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	Unclear risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? Unclear (37.5% of studies not described whether the reference test was interpreted without knowledge of the FDG-PET findings) Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Low risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to

		flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? Yes Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	Low risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

PS - Staging (bone marrow disease)

Author, year	Technology	Number of participants	Population	Sensitivity %	Specificity %
Pelosi 2011	FDG-PET/CT	207	patients at initial staging with non-Hodgkin's lymphoma	64.4	100
	bone marrow biopsy			67.8	100

PS - Pelosi 2011

Clinical features and settings	Hodgkin's lymphoma, aggressive non Hodgkin's; Country: Italy
Participants	337 consecutive patients with Hodgkin's lymphoma (130) or aggressive non Hodgkin's lymphoma (207; diffuse large B-cell 120, follicular grade II-III 48, mantle cell 7, others 32). Age (yr) median 49.4, range 11–84; male 189, female 148
Study design	prospective diagnostic accuracy study, with consecutive accrual
Target condition and reference standard(s)	Bone marrow extension of disease Reference standard: composite: bone marrow biopsy and imaging follow up (with FDG-PET/CT and /or MRI)
Index and comparator tests	FDG-PET/CT comparator: bone marrow biopsy
Follow-up	not reported
Notes	data reported also for disease subgroup

Assessment of methodological quality table

Item	Authors' judgement	Support for judgement
1A. PATIENT SELECTION (risk of bias)	Low risk	Describe methods of 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 Could the selection of patients have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
2A. INDEX TEST(S) (risk of bias)	High risk	Describe the index test and how it was conducted and interpreted: Were the index test results interpreted without knowledge of the results of the reference standard? No (incorporation bias)

		If a threshold was used, was it pre-specified? Yes Could the conduct or interpretation of the index test have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
3A. REFERENCE STANDARD (risk of bias)	High risk	Describe the reference standard and how it was conducted and interpreted: Is the reference standard likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index test? No (incorporation bias) Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
4. FLOW AND TIMING (risk of bias)	Unclear risk	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard: Was there an appropriate interval between index test(s) and reference standard? Yes Did all patients receive a reference standard? Yes Did patients receive the same reference standard? No Were all patients included in the analysis? Yes Could the patient flow have introduced bias? RISK: LOW=YES/HIGH=NO/UNCLEAR
1.B PATIENT SELECTION (concern of applicability)	Low risk	Describe included patients (prior testing, presentation, intended use of index test and setting): Is there concern that the included patients do not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
2.B INDEX TEST(S) (concern of applicability)	Low risk	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR
3B. REFERENCE STANDARD (concern of applicability)	High risk	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: LOW=YES/HIGH=NO/UNCLEAR

Characteristics of excluded studies

Abdulqadhr 2011

Reason for exclusion	not available diagnostic accuracy estimates (sensitivity, specificity)

Alavi 2011

Reason for exclusion	retrospective study

Cheng 2011

Reason for exclusion	retrospective study

Fujiwara 2011

Reason for exclusion	retrospective study

Gu 2011

Reason for exclusion	FDG-PET as reference standard, lesion-based analysis

Hong 2011

Reason for exclusion	retrospective study

Huang 2011

Reason for exclusion	retrospective study

Ilica 2011

Reason for exclusion	narrative review

Mittal 2011

Reason for exclusion	unclear design of study

Papajik 2011

Reason for exclusion	FDG-PET/CT as reference standard

Rodriguez-Vigil 2011

Reason for exclusion	agreement study design

van Ufford 2011

Reason for exclusion	agreement study

Search strategies

1 CDSR, DARE, HTA database, CENTRAL search strategy

1. “Positron-Emission Tomography” [MeSH descriptor explode all trees]
2. “Fluorodeoxyglucose F18” [MeSH descriptor explode all trees]
3. "positron emission tomography":ti,ab,kw
4. pet*: ti,ab,kw
5. pet scan*: ti,ab,kw
6. "Fluorodeoxyglucose F18": ti,ab,kw
7. fdg NEAR/2 18: ti,ab,kw
8. 1/7 OR
9. “Lymphoma”/exp
10. 8 AND 9

Publication date: January 2011 - December 2011

2 MEDLINE search strategy

1. "Fluorodeoxyglucose F18"[Mesh]
2. "2-Fluoro-2-deoxyglucose" [All Fields]
3. "18F Fluorodeoxyglucose" [All Fields]
4. "F 18 Fluorodeoxyglucose" [All Fields]
5. Fludeoxyglucose* [All Fields]
6. "2 fluoro 2 deoxy d glucose"[All Fields]
7. 18fluorodesoxyglucose*[All Fields]
8. fluorodeoxyglucose*[All Fields]
9. "fluorine 18 fluorodeoxyglucose" [All Fields]
10. 18f dg*[All Fields])
11. 18fluorodeoxyglucose*[All Fields]
12. 18fdg [All Fields]
13. 18 fdg* [All Fields]
14. fdg 18* [All Fields]
15. fdg/* [All Fields]
16. "fdg pet"[All Fields]
17. "Positron-Emission Tomography"[Mesh]
18. "positron emission tomography" [title/abstract]
19. pet [title/abstract]
20. "pet scan" [All Fields]
21. "pet scans" [All Fields]
22. "pet scanner" [All Fields]
23. petscan [All Fields]
24. 1/23 OR
25. "Lymphoma"[Mesh:noexp]
26. "Hodgkin Disease"[Mesh]
27. "Lymphoma, Non-Hodgkin"[Mesh]
28. "lymphomas"[Title/Abstract]
29. "lymphoma"[Title/Abstract]
30. "hodgkin's"[Title/Abstract]
31. "hodgkins"[Title/Abstract]
32. "hodgkin"[Title/Abstract]
33. "hodgkin s"[Title/Abstract]
34. "lymphogranuloma"[Title/Abstract]
35. "non hodgkin"[Title/Abstract]
36. "non hodgkin s b"[Title/Abstract]
37. "non hodgkin's"[Title/Abstract]
38. "reticulum cell sarcoma"[Title/Abstract]
39. "reticulum cell sarcomas"[Title/Abstract]
40. "reticulosarcoma"[Title/Abstract]
41. "reticulosarcomas"[Title/Abstract]
42. "lymphosarcoma"[Title/Abstract]

43. "lymphosarcomas"[Title/Abstract]
44. "lymphatic sarcoma"[Title/Abstract]
45. "lymphatic sarcomas"[Title/Abstract]
46. "burkitt's"[Title/Abstract]
47. "burkitt"[Title/Abstract]
48. "burkitt s"[Title/Abstract]
49. "lymphocytic leukemia"[Title/Abstract]
50. "lymphocytic leukemias"[Title/Abstract]
51. "lymphomatoid granulomatoses"[Title/Abstract]
52. "lymphomatoid granulomatosis"[Title/Abstract]
53. "brill symmers disease"[Title/Abstract]
54. "immunoblastoma"[Title/Abstract]
55. "immunoblastomas"[Title/Abstract]
56. "immunoblastosarcoma"[Title/Abstract]
57. "immunoblastosarcomas"[Title/Abstract]
58. "immunoblastic sarcoma"[Title/Abstract]
59. "immunoblastic sarcomas"[Title/Abstract]
60. "granulomatous slack skin"[Title/Abstract]
61. "lymphomatoid papulosis"[Title/Abstract]
62. "mycosis fungoides"[Title/Abstract]
63. "pagetoid reticulosis"[Title/Abstract]
64. "woringer kolopp disease"[Title/Abstract]
65. "ketron goodman disease"[Title/Abstract]
66. "sezary lymphoma"[Title/Abstract]
67. "sezary's syndrome"[Title/Abstract]
68. "sezary syndrome"[Title/Abstract]
69. 25/68 OR
70. 69 AND 24
71. "editorial"[Publication Type]
72. "comment"[Publication Type]
73. "letter"[Publication Type]
74. 71/73 OR
75. 70 NOT 74

LIMITS: Humans
Languages: English, French, Italian, Spanish
Publication date: January 2011 - December 2011

3 EMBASE search strategy

1. 'positron emission tomography'/syn
2. 'fluorodeoxyglucose f 18'/exp
3. 'fluorodeoxyglucose f 18'/syn
4. 'computer assisted emission tomography'/exp
5. 'computer assisted emission tomography' OR
6. pet
7. 'pet scans'
8. 'pet scanner'
9. ‘pet scan'
10. 'pet/ct scan'
11. 'pet/ct scans'
12. 'pet/ct'
13. 'positron emission tomography/computed tomography'
14. pet NEAR/4 scan*
15. pet NEAR/4 ct
16. 1/23 OR
17. 'lymphoma'/de
18. 'hodgkin disease'/exp
19. 'classical hodgkin lymphoma'/exp
20. 'nonhodgkin lymphoma'/exp
21. 'intestine lymphoma'/de
22. 'skin lymphoma'/exp
23. 'reed sternberg cell'/de

24. 'cutaneous t cell lymphoma'/exp
25. 'histiocytic lymphoma'/exp
26. 'marginal zone lymphoma'/exp
27. 't cell lymphoma'/exp
28. 17/27 OR
29. 16 AND 28
LIMITS:
Publication Type: article; article in press; erratum; short survey
Humans
Languages: English, French, Italian, Spanish
Publication date: January 2011 - December 2011

Appendix 21

Search strategy systematic review of economic evaluation

	Studi economici strategia di ricerca Cochrane Library					
"Positron-Emission Tomography" [MeSH descriptor explode all trees] OR "Fluorodeoxyglucose F18" [MeSH descriptor explode all trees] OR. "positron emission tomography": ti,ab,kw OR. pet*: ti,ab,kw OR. pet scan*: ti,ab,kw OR. "Fluorodeoxyglucose F18": ti,ab,kw OR. fdg NEAR/2 18: ti,ab,kw AND CT OR "computer tomography"	AND	**Staging**	AND	**Neoplasms "[Mesh explodes all trees]** OR Neoplasm*: ti,ab,kw OR Cancer* : ti,ab,kw OR Tumor*" : ti,ab,kw OR Tumour*: ti,ab,kw OR Carcinoma: ti,ab,kw	**And**	economic AND (evaluation OR analysis or assessment) OR "cost effectiveness": ti,ab,kw OR "cost-effectiveness": ti,ab,kw OR CEA: ti,ab,kw OR "cost utility" : ti,ab,kw OR "cost-utility" : ti,ab,kw OR CUA : ti,ab,kw OR "cost benefit": ti,ab,kw OR "cost-benefit" : ti,ab,kw OR CBA : ti,ab,kw OR "cost analysis" : ti,ab,kw OR "cost consequence" : ti,ab,kw OR "cost-consequence" : ti,ab,kw OR "cost minimization" : ti,ab,kw OR "cost-minimization" : ti,ab,kw

Limiti: 2011-2012

Embase						
studi economici						
"positron emission tomography"/syn OR "fluorodeoxyglucose f 18"/exp OR "fluorodeoxyglucose f 18"/syn OR "computer assisted emission tomography"/exp OR "computer assisted emission tomography" OR pet OR "pet scans" OR "pet scanner" OR "pet scan" OR "pet/ct scan" OR "pet/ct scans" OR "pet/ct" OR OR"positron emission tomography/computed tomography" OR OR pet NEAR/4 scan* OR pet NEAR/4 ct AND CT OR "computer tomography"	AND	Staging	AND	"Prostatic Neoplasms"/syn OR "prostatic cancer" OR "prostatic neoplasm" OR "prostate neoplasms" OR "prostate cancer" OR "Cancer of the Prostate" OR Prostatic NEAR/4 cancer* OR Prostatic NEAR/4 neoplasm	AND	economic AND (evaluation OR analysis or assessment) OR "cost effectiveness" OR "cost-effectiveness" OR CEA OR "cost utility" OR "cost-utility" OR CUA OR "cost benefit" OR "cost-benefit" OR CBA OR "cost analysis" OR "cost consequence" OR "cost-consequence" OR "cost minimization" OR "cost-minimization"
Limiti: da gennaio 2011; humans "article" OR "review"/it OR "short survey"						

Strategia studi economici / Medline						
"Fluorodeoxyglucose F18"[Mesh] **OR** "2-Fluoro-2-deoxyglucose" [All Fields] **OR** "18F Fluorodeoxyglucose" [All Fields] **OR** "F 18 Fluorodeoxyglucose" [All Fields] **OR** Fludeoxyglucose* [All Fields] **OR** "2 fluoro 2 deoxy d glucose"[All Fields] **OR** 18fluorodesoxyglucose*[All Fields] **OR** fluorodeoxyglucose*[All Fields] **OR** "fluorine 18 fluorodeoxyglucose" [All Fields] **OR** 18f dg*[All Fields]) **OR** 18fluorodeoxyglucose*[All Fields] **OR** 18fdg [All Fields] **OR** 18 fdg* [All Fields] **OR** fdg 18* [All Fields] **OR** fdg/* [All Fields] **OR** "fdg pet"[All Fields] **OR** "Positron-Emission Tomography"[Mesh] **OR** "positron emission tomography" [title/abstract] **OR** pet [title/abstract] **OR** "pet scan" [All Fields] **OR** "pet scans" [All Fields] **OR** "pet scanner" [All Fields] **OR** petscan [All Fields] AND CT OR "computer tomography"	AND	**Staging**	AND	Neoplasms **[Mesh explodes all trees] contiene** OR Neoplasm* [Title/Abstract] OR Cancer* [Title/Abstract] OR Tumor* [Title/Abstract] OR Carcinoma* [Title/Abstract] OR Tumour* [Title/Abstract] OR Oncology [Title/Abstract]	**And**	economic AND (evaluation OR analysis or assessment) " [All Fields] OR "cost effectiveness" " [All Fields] OR "cost-effectiveness" " [All Fields] OR CEA " [All Fields] OR "cost utility" " [All Fields] OR "cost-utility" " [All Fields] OR CUA " [All Fields] OR "cost benefit" " [All Fields] OR "cost-benefit" " [All Fields] OR CBA " [All Fields] OR "cost analysis" " [All Fields] OR "cost consequence" " [All Fields] OR "cost-consequence" " [All Fields] OR "cost minimization" " [All Fields] OR "cost-minimization" " [All Fields]

Limiti: 2011 – 2012

Humans