La tomographie par émission de positrons en Belgique: une mise à jour

KCE reports 110B

Federaal Kenniscentrum voor de Gezondheidszorg
Centre fédéral d'expertise des soins de santé
2009

Le Centre fédéral d'expertise des soins de santé

Présentation : Le Centre fédéral d'expertise des soins de santé est un parastatal, créé le 24 décembre 2002 par la loi-programme (articles 262 à 266), sous tutelle du Ministre de la Santé publique et des Affaires sociales, qui est chargé de réaliser des études éclairant la décision politique dans le domaine des soins de santé et de l'assurance maladie.

Conseil d'administration

Membres effectifs : Gillet Pierre (Président), Cuypers Dirk (Vice-président), Avontroodt Yolande, De Cock Jo (Vice -président), Baeyens Jean-Pierre, De Ridder Henri, de Stexhe Olivier, Godin Jean-Noël, Goyens Floris, Maes Jef, Mertens Pascal, Mertens Raf, Moens Marc, Perl François, Van Massenhove Frank (Vice-président), Degadt Peter, Verertbruggen Patrick, Schetgen Marco, Devos Daniël, Smeets Yves.

Membres suppléants : Cuypers Rita, Decoster Christiaan, Collin Benoit, Stamatakis Lambert, Vermeyen Karel, Kesteloot Katrien, Ooghe Bart, Lernoux Frederic, Vanderstappen Anne, Palsterman Paul, Messiaen Geert, Remacle Anne, Lemye Roland, Poncé Annick, Smiets Pierre, Bertels Jan, Lucet Catherine.

Commissaire du gouvernement : Roger Yves

Direction

Directeur général a.i. : Jean-Pierre Closon

Contact

Centre fédéral d'expertise des soins de santé (KCE).
Cité Administrative Botanique, Doorbuilding (10ème)
Boulevard du Jardin Botanique, 55
B-1000 Bruxelles
Belgium

Tel: +32 [0]2 287 33 88
Fax: +32 [0]2 287 33 85

Email : info@kce.fgov.be
Web : http://www.kce.fgov.be

La tomographie par émission de positrons en Belgique : une mise à jour

KCE reports vol 110B

JOAN VLAYEN, SABINE STORDEUR, ANN VAN DEN BRUEL,
FRANÇOISE MAMBOURG, MARIJKE EYSSEN

Federaal Kenniscentrum voor de gezondheidszorg
Centre fédéral d'expertise des soins de santé
2009

KCE reports 110B

Titre:	La tomographie par émission de positrons en Belgique: une mise à jour
Auteurs:	Joan Vlayen, Sabine Stordeur, Ann Van Den Bruel, Françoise Mambourg, Marijke Eyssen
Experts externes:	M. Peeters (UZ Gent)[1,2], P. Flamen (Institut Bordet, Bruxelles)[1,2], E. Danse (Cliniques universitaires Saint-Luc, Bruxelles)[1,2], J-L Van Laethem (ULB, Brussel)[2], R. Hustinx (ULg, Liège)[2], D. Galdermans (ZNA, Antwerpen)[1,2], M. Lonneux (Cliniques universitaires Saint-Luc, Bruxelles)[1,2], M. Lemort (Institut Bordet, Bruxelles)[1,2], J-F Baurain (Cliniques universitaires Saint-Luc, Bruxelles)[2], S. Goldman (ULB, Bruxelles)[1], M. Van Goethem (UZA, Antwerpen)[1,2], G. Villeirs (UGent)[1,2], J. Roland (ZNA, Antwerpen)[1,2], F. Kridelka (ULg, Liège)[1,2], B. Tombal (Cliniques universitaires Saint-Luc, Bruxelles)[2], G. Jerusalem (ULg, Liège)[2], K. Haustermans (UZ Leuven)[2], I. Vergote (UZ Leuven)[2], W. Van Paesschen (UZ Leuven)[1,2], A. Maes (AZ Groeninge, Kortrijk)[1,2], L. Mesotten (ZOL, Genk)[1,2], O. De Winter (OLV Aalst)[1,2], E. Salmon (ULg, Liège)[2], F. Jacobs (ULB, Bruxelles)[2]. [1] présent à la réunion d'experts; [2] a formulé des remarques
Remerciements:	Irina Cleemput (KCE), Hans Van Brabandt (KCE), Kris Henau (Fondation Registre du cancer)
Validateurs externes:	Sigrid Stroobants (UZ Antwerpen), Ahmad Hussein Awada (Institut Bordet), Elizabeth Adams (VATAP, USA)
Conflits d'intérêt:	Tous les experts exercent une activité dans un hôpital disposant d'un scanner PET.
Disclaimer :	Les experts externes ont été consultés sur une version (préliminaire) du rapport scientifique. Une version (finale) a ensuite été soumise aux validateurs. La validation du rapport résulte d'un consensus ou d'un vote majoritaire entre les validateurs. Le KCE reste seul responsable des erreurs ou omissions qui pourraient subsister de même que des recommandations faites aux autorités publiques.
Mise en page :	Ine Verhulst

Bruxelles, 29 juin 2009

Etude nr 2009-51

Domaine : Health Technology Assessment

MeSH : Positron-Emission Tomography; Fluorodeoxyglucose F18; Technology Assessment, Biomedical; Sensitivity and Specificity

NLM classification : WN206

Language : français, anglais

Format : Adobe® PDF™ (A4)

Dépôt légal : D/2009/10.273/25

Vlayen J, Stordeur S, Van den Bruel A, Mambourg F, Eyssen M. La tomographie par émission de positrons en Belgique: une mise à jour. Health Technology Assessment (HTA). Bruxelles: Centre fédéral d'expertise des soins de santé (KCE); 2009. KCE reports 110B (D/2009/10.273/25)

PREFACE

On se souviendra que le KCE avait publié un premier rapport sur la tomographie par émission de positrons en 2005. Ce rapport n'avait pas fait que des heureux dans la mesure où il concluait que 10 PET scanners seulement étaient nécessaires pour couvrir les besoins des patients belges. En effet, à l'époque, les cas où l'efficacité de cette technique était vraiment démontrée étaient encore relativement rares.

Sur base des conclusions de l'étude du KCE, les autorités de santé avaient établi une programmation de 13 appareils au maximum. Celle-ci doit cependant être revue pour deux raisons. D'une part une plainte a été introduite auprès de la Commission européenne contre les critères de programmation utilisés, obligeant la Belgique à les revoir pour les rendre plus objectifs. D'autre part, la technique d'imagerie est en train d'évoluer du PET au PET/CT et la littérature scientifique des quatre dernières années a mis en évidence de nouvelles indications intéressantes, particulièrement dans la prise en charge de certains cancers.

L'ensemble de ces éléments a amené Madame la Ministre au début de cette année, à inviter le KCE à actualiser d'urgence son rapport de 2005 de façon à mettre sur pied une programmation qui tienne compte à la fois des impératifs européens et de l'évolution de la science.

Nous remercions tous les experts qui ont participé à nos réunions de travail préalables au dépôt de nos conclusions et espérons que ces dialogues constructifs permettront d'évoluer vers une solution qui rencontre l'intérêt et les préoccupations de toutes les parties.

Jean-Pierre CLOSON
Directeur général a.i.

Résumé

INTRODUCTION

Il y a 4 ans environ, le KCE a publié un premier rapport sur la tomographie par émission de positrons (PET), une technologie de diagnostic non invasive qui permet de visualiser une anomalie métabolique au niveau des organes ou des tissus atteints dans certaines pathologies bien déterminées (essentiellement en oncologie). A cette fin, le PET utilise un radio-isotope (traceur) qui est injecté au patient. Le traceur le plus utilisé est le 18 Fluorodesoxy-glucose (FDG).

Une évolution récente a conduit à l'utilisation d'appareils PET/CT et permet désormais de combiner en une seule image les résultats du PET et du CT scanner. L'avantage de cette image unique est qu'elle associe des informations anatomiques (CT) et fonctionnelles (PET).

En 2005, le nombre de scanners PET nécessaires en Belgique avait été estimé à 10. Réglementairement, le nombre maximal de PET autorisés a été fixé à 13. Toutefois, sur base du nombre d'examens PET facturés, on peut supposer que le parc actuel compte un nombre d'appareils supérieur à 13.

À l'heure actuelle, la programmation des scanners PET est réglementée par la loi sur les hôpitaux du 7 août 1987, la loi du 27 avril 2005 et plusieurs arrêtés royaux. Récemment, la Commission européenne a été saisie d'une plainte contre cette règlementation motivée par l'argument selon lequel les règles de programmation belges des scanners PET seraient fondées sur des critères non objectifs. La Commission européenne a accepté de classer la plainte sans suite, à condition que la Belgique revoie sa programmation et la fonde exclusivement sur des critères objectifs.

Le présent rapport a principalement pour objectif de mettre à jour les indications cliniques du PET. Par ailleurs, il fournit un aperçu des critères utilisés pour la programmation du PET dans d'autres pays. Enfin, l'estimation du nombre de patients ayant besoin d'un PET en Belgique est également débattue.

METHODOLOGIE UTILISEE POUR LA REVUE DE LA LITTÉRATURE

Une revue systématique de la littérature a été conduite pour évaluer la précision diagnostique et l'efficacité clinique du PET et du PET/CT. Dans un premier temps, nous nous sommes intéressés aux rapports HTA, aux revues systématiques de la littérature et aux méta-analyses, publiés depuis le rapport précédent du KCE sur le PET (2005).

Pour chaque indication, nous avons ensuite recherché les études primaires, incluant les essais cliniques randomisés et les études diagnostiques et pronostiques.

L'efficacité du PET et du PET/CT a été évaluée en oncologie, en cardiologie, en neurologie et en infectiologie. Pour quelques indications particulières (essentiellement neurologiques), comme la démence, les tumeurs cérébrales et la maladie de Parkinson, nous avons également recherché des études ayant utilisé d'autres traceurs que le FDG. Un seul expert du KCE s'est chargé de la sélection méthodique des études. La qualité des études a été appréciée au moyen de grilles d'évaluation standardisées.

Nous avons distingué quatre niveaux d'efficacité diagnostique : (1) précision technique (2) précision diagnostique (3) impact sur le résultat pour le patient et (4) rapport coût-efficacité.

Comme d'habitude, les conclusions de la revue de la littérature ont été soumises à une équipe multidisciplinaire d'experts externes et débattues au cours de 4 réunions d'experts distinctes, une par discipline médicale. Le rapport final a été validé par trois validateurs externes.

INDICATIONS CLINIQUES DU PET ET DU PET/CT

Les indications oncologiques du PET et du PET/CT restent les mieux étudiées. Plusieurs indications sont venues s'ajouter à la liste actuelle des indications remboursées. Certaines de ces indications étaient étayées par des preuves solides. Pour d'autres, le PET et le PET/CT sont sans doute utiles, mais les preuves scientifiques sont encore insuffisantes (voir tableau 1). Il est intéressant de noter que plusieurs indications pour lesquelles le PET est actuellement remboursé ne sont pas ou insuffisamment étayées par des preuves scientifiques. À l'inverse, il y a également des indications pour lesquelles des éléments probants ont été trouvés en suffisance, mais qui ne sont pas encore remboursées (voir tableau 1).

À ce jour, les tumeurs suivantes ne constituent pas une indication pour le PET et le PET/CT et n'ont donc pas été reprises dans le tableau 1 : cancer primitif du foie, cancer de l'estomac, cancer du sein, cancer des testicules, cancer de la vessie, cancer de la prostate, cancer de l'utérus et cancer du pénis.

Il est important de mentionner que le niveau de preuve ne s'est pas amélioré au cours des 4 dernières années. De nombreuses études inclues sont limitées par la petite taille de la population investiguée, une approche rétrospective et de nombreuses formes de biais.

Tableau 1. Indications potentielles pour le PET(/CT).

Indication	Remboursement actuel	Niveau de preuve #
Cancer du poumon		
Appréciation d'un nodule pulmonaire solitaire	Oui	Oui (2)
Stadification initiale du carcinome pulmonaire non à petites cellules	Oui	Oui (4)
Planification de la radiothérapie	Non	Sans conclusion (2)
Appréciation de la masse résiduelle ou de la récidive dans le carcinome pulmonaire non à petites cellules	Oui	Sans conclusion (2)
Lymphome		
Stadification initiale du lymphome de Hodgkin ou du lymphome non Hodgkinien (stade intermédiaire ou avancé)	Oui	Oui (2)
Appréciation de la masse résiduelle ou de la récidive d'un lymphome	Oui	Oui pour la masse résiduelle (2) Sans conclusion pour la récidive du lymphome (2)
Tumeurs de la tête et du cou		
Stadification initiale	Non	Oui (2)
Appréciation de la masse résiduelle ou de la récidive des tumeurs de la bouche ou du pharynx	Oui	Oui (2)
Carcinome avec tumeur primitive inconnue		
Dépistage de la tumeur primitive	Non	Oui (2)
Cancer colorectal		
Évaluation préopératoire de métastases au foie potentiellement opérables d'un cancer colorectal	Non	Oui (2)
Appréciation de la masse résiduelle ou de la récidive d'un cancer colorectal	Oui	Oui (3)
Mélanome malin		
Stadification initiale du mélanome malin (stade IIc ou supérieur)	Oui	Oui (2)

Indication	Remboursement actuel	Niveau de preuve [#]
Appréciation de la masse résiduelle ou de la récidive d'un mélanome malin	Oui	Sans conclusion pour la masse résiduelle et détection d'une récidive (2) Oui pour la stadification d'une récidive (2)
Cancer de l'œsophage		
Stadification initiale du cancer de l'oesophage	Oui	Oui (2)
Suivi de l'effet du traitement	Non	Sans conclusion (2)
Cancer de la thyroïde		
Appréciation des nodules thyroïdiens dont les résultats cytologiques ne sont pas concluants	Non	Sans conclusion (2)
Cancer du pancréas		
Distinction entre une pancréatite chronique et le cancer du pancréas, de même qu'entre les kystes pancréatiques bénins et malins	Non	Oui (2)
Stadification initiale du cancer du pancréas	Oui	Sans conclusion (2)
Appréciation de la masse résiduelle ou de la récidive d'un cancer du pancréas	Oui	Non
Cancer du col de l'utérus		
Stadification initiale	Non	Oui (2)
Appréciation de la récidive	Non	Sans conclusion (2)
Cancer des ovaires		
Diagnostic initial	Non	Sans conclusion (2)
Appréciation de la masse résiduelle ou de la récidive d'un cancer des ovaires	Oui	Oui (2)
GIST/TSGI (tumeurs stromales gastro-intestinales)		
Suivi de l'effet du traitement	Non	Oui (2)
Cancer du cerveau		
Appréciation de la masse résiduelle ou de la récidive d'un cancer du cerveau	Oui	Sans conclusion (2)
Cardiologie		
Appréciation de la viabilité du myocarde	Oui	Oui (2)
Neurologie		
Diagnostic de la maladie d'Alzheimer chez les patients souffrant de démence	Non	Sans conclusion (2)
Appréciation préopératoire de l'épilepsie réfractaire	Oui	Sans conclusion (2)
Pathologies infectieuses		
Appréciation de l'ostéomyélite chronique	Non	Sans conclusion (2)
Appréciation des infections associées à des prothèses	Non	Sans conclusion (2)
Appréciation d'une fièvre d'origine inconnue	Non	Sans conclusion (2)

[#] Comme nous l'avons dit plus haut, nous faisons la distinction entre les niveaux suivants d'efficacité diagnostique : (1) précision technique (2) précision diagnostique (3) impact sur le résultat pour le patient et (4) rapport coût-efficacité.

PROGRAMMATION DES SCANNERS PET

À l'heure actuelle, le nombre maximum de scanners PET est limité dans notre pays sur base des critères suivants : 1 appareil par hôpital universitaire (n = 7), 1 appareil par hôpital offrant des soins chirurgicaux et médicaux exclusivement dans le domaine de l'oncologie (n = 1) et 1 appareil pour 1.6 million d'habitants (n = 5 ; 3 en Flandre et 2 en Wallonie). Au total, le parc installé comprend dès lors 13 scanners. Pour répartir ces 13 appareils, les régions et les communautés appliquent les normes d'agrément spécifiques suivantes : la preuve d'une activité oncologique suffisante ; la présence d'une caméra gamma ; la disponibilité d'un personnel médical comprenant au moins 3 spécialistes agréés à temps plein en médecine nucléaire, un physicien ou un ingénieur à temps plein et 2 infirmiers à temps plein travaillant exclusivement dans le service ; un enregistrement interne et un contrôle de la qualité externe.

A l'heure actuelle, le remboursement du PET se limite à 16 indications (voir le tableau 1). Cela étant, le remboursement d'indications officiellement non remboursées est possible via le code de nomenclature « double tomographie ». En 2007, quelque 18.500 scanners PET officiels ont été remboursés (~3 millions d'euros) et on estime le nombre des remboursements non officiels à 20.000 (~5,5 millions d'euros).

UTILISATION DE LA PROGRAMMATION, DES CRITÈRES D'AGRÉMENT ET DES MODALITÉS DE REMBOURSEMENT DANS D'AUTRES PAYS

Quelque 14 agences HTA de 11 pays différents ont répondu à une question récente de l'International Network of Agencies for Health Technology Assessment (INAHTA) sur les critères utilisés pour déterminer le nombre de scanners PET dans chaque pays. Dans le présent rapport du KCE, ces résultats ont été complétés par une recherche Internet pour obtenir des informations supplémentaires.

Certains pays, comme la France et Israël, programment le nombre de scanners PET en fonction de la taille de la population. Dans certains pays, comme la France et l'Australie, il existe aussi des normes d'agrément similaires à celles utilisées en Belgique. Enfin, certains pays, comme l'Australie, l'Espagne et les États-Unis, limitent le nombre d'indications remboursées. Pour autant que nous sachions, la Belgique est le seul pays qui combine l'ensemble de ces critères et modalités, et qui programme le nombre de scanners PET sur la base d'autres critères que celui de la population.

NOMBRE DE PATIENTS AYANT BESOIN D'UN EXAMEN PET

En principe, on peut distinguer deux méthodes de calcul du nombre de patients ayant besoin d'un examen PET. L'approche prospective, qui part uniquement des indications basées sur des preuves et des données épidémiologiques, présente l'avantage d'être complète. Le gros problème est que ces données épidémiologiques ne sont ni disponibles en suffisance ni assez détaillées.

Une approche rétrospective (basée sur l'utilisation réelle des scanners PET), telle que celle utilisée dans le précédent rapport du KCE, offre l'avantage de partir de données enregistrées (obligatoires). En conséquence, la somme des indications pour lesquelles un examen PET est préconisé devrait refléter correctement le nombre total de patients. Un inconvénient important de cette approche est que les patients qui ont besoin d'un examen PET, mais qui n'en bénéficient pas ne sont pas comptabilisés. Avec pour conséquence, une sous-estimation du besoin réel. En outre, les données ne sont pas librement accessibles. Pour ces raisons, contrairement au rapport précédent, ces données n'ont pas été utilisées pour ce rapport-ci.

En d'autres termes, calculer les besoins d'examens PET est soit impossible pour le moment (approche prospective), soit inadéquat (approche rétrospective). En outre les indications fondées sur des preuves sont en constante évolution, ce qui rend difficile la mise à jour permanente de l'évaluation des besoins.

CONCLUSIONS

Quatre constats importants ressortent du présent rapport :

1. Depuis 4 ans, on a observé une augmentation des preuves scientifiques pour de nouvelles indications qui, pour l'instant, ne sont pas encore remboursées.
2. En autorisant un remboursement des examens PET via le code de la nomenclature « double tomographie », la programmation n'a qu'un impact limité sur le nombre réel d'examens PET.
3. Il existe deux méthodes pour mettre en adéquation le nombre de scanners PET et les besoins : la programmation, d'une part, et les normes d'agrément et modalités de remboursement, d'autre part.
4. En Belgique, la programmation des scanners PET fondée sur une estimation des besoins n'est pas possible à court terme.

RECOMMANDATIONS

- Une programmation des scanners PET sur la base d'une estimation des besoins n'est pas possible à court terme et n'est dès lors pas recommandée.
- Une alternative à la programmation est de réguler le nombre de scanners PET en:
    - fixant des normes d'agrément suffisamment sévères et en veillant à leur application stricte pour garantir la qualité des prestations;
    - déterminant des conditions de remboursement qui limitent les prestations facturables à celles qui correspondent à des indications fondées sur des preuves scientifiques.
- Le remboursement de ces prestations doit être lié à l'enregistrement préalable de l'indication dans un registre unique, informatisé et standardisé. Ce registre permettra en outre de vérifier en permanence que le système proposé répond aux besoins actuels.
- La liste limitative d'indications remboursées doit être revue tous les 3 ans, en accordant une attention particulière aux nouveaux traceurs et aux nouvelles techniques d'imagerie. A l'occasion de cette révision trisannuelle, la question de recherche doit être élargie aux autres techniques diagnostiques, afin de positionner de manière systématique le PET et le PET/CT en regard de ces techniques.
- Pour les indications oncologiques qui se sont ajoutées à la liste et pour lesquelles il n'y a pas encore de preuves scientifiques concluantes, un remboursement éventuel devrait être conditionné par la tenue d'une consultation oncologique multidisciplinaire.
- Le respect des critères de remboursement doit être contrôlé de manière systématique.
- Le remboursement des examens PET via le code de nomenclature « double tomographie » doit être supprimé afin de pouvoir suivre l'évolution du nombre d'examens de manière transparente et contrôlée.

Scientific summary

Table of contents

LIST OF ABBREVIATIONS

1.5-T	1.5 Tesla
95%CI	95% confidence interval
AATRM	Catalan Agency for Health Technology Assessment
ACCC	Association of Comprehensive Cancer Centres
ACCP	American College of Chest Physicians
AD	Alzheimer's disease
AETSA	Agencia de Evaluación de Tecnologias Sanitarias de Andalucia
AHRQ	Agency for Healthcare Research and Quality
AHTAPol	Agency for Health Technology Assessment in Poland
ALND	Axillary lymph node dissection
AUC	Area under the curve
AUS	Axillary ultrasonography
BCBS	Blue Cross-Blue Shield
Bq	Becquerel
CADTH	Canadian Agency for Drugs and Technologies in Health
CEA	Carcino embryonic antigen
CCOHTA	Canadian Coordinating Office for Health Technology Assessment
CMS	Centers for Medicare and Medicaid Services
CRD	Centre for Reviews and Dissemination
CT	Computed tomography
CUP	Carcinoma of unknown primary
DACEHTA	Danish Centre for Evaluation and Health Technology Assessment
DAT	Dopamine transporters
DFS	Disease-free survival
DIMDI	German Institute of Medical Documentation and Information
ECRI	Emergency Care Research Institute
EEG	Electro-encephalogram
EFNS	European Federation of Neurological Societies
ERCP	Endoscopic Retrograde Cholangiopancreatography
EUS	Endoscopic ultrasonography
FDG	Fluoro-deoxyglucose
FET	Fluoro-ethyl L-thyrosine
FIGO	Fédération Internationale de Gynécologie et d'Obstétrique
FLT	Fluoro-L-thymidine
FN	False negatives
FNAC	Fine-needle aspiration cytology
FNCLCC	Fédération Nationale des Centres de Lutte Contre le Cancer
FP	False positives

FUO	Fever of unknown origin
g	gram
GIST	Gastrointestinal Stromal Tumour
HAS	Haute Authorité de Santé
HL	Hodgkin lymphoma
HNSCC	Head and neck squamous cell cancer
IQWIG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
INAHTA	International Network of Agencies for Health Technology
HNSCC	Head and neck squamous cell cancer
HR	Hazard ratio
HTA	Health technology assessment
ILN	Inguinal lymph nodes
KCE	Belgian Healthcare Knowledge Centre
LABC	Locally-advanced breast cancer
LR	Likelihood ratio
MAS	Medical Advisory Secretariat
MDCT	Multidetector computed tomography
MET	Methionine
MLN	Mediastinal lymph nodes
MRI	Magnetic resonance imaging
MRS	Magnetic resonance spectroscopy
MSAC	Medical Services Advisory Committee
M-staging	Metastasis staging
NCCHTA	National Coordinating Centre for Health Technology Assessment
NHL	Non-Hodgkin lymphoma
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NSCLC	Non-small-cell lung cancer
N-staging	Nodal staging
OR	Odds ratio
OS	Overall survival
PALN	Para-aortic lymph nodes
pCR	Pathological complete remission
PD	Parkinson disease
PET	Positron emission tomography
PFS	Progression-free survival
PLN	Pelvic lymph nodes
QALY	Quality-adjusted life year
RCC	Renal cell cancer
RCT	Randomised controlled trial

RIZIV/INAMI	Rijksinstituut voor Ziekte- en Invaliditeitsverzekering/Institut National d'Insurance Maladie et Invalidité
RNA	Ribonucleic acid
RR	Relative risk
RT	Radiotherapy
SBU	Swedish Council on Technology Assessment in Health Care
SCC	Squamous cell cancer
SCLC	Small-cell lung cancer
SCM	Scintimammography
Se	Sensitivity
SIGN	Scottish Intercollegiate Guidelines Network
SLNB	Sentinel lymph node biopsy
Sp	Specificity
SPECT	Single photon emission computed tomography
SPN	Solitary pulmonary nodule
SR	Systematic review
SROC	Summary receiver operating characteristics
Tg	Thyroglobulin
TN	True negatives
TP	True positives
TRUS	Transrectal ultrasonography
TSH	Thyroid-stimulating hormone
TTP	Time-to-progression
SUV	Standardised uptake value
US	Ultrasonography
USA	United States of America
WBS	White blood cell scinitgraphy
ZonMw	Nederlandse organisatie voor gezondheidsonderzoek en zorginnovatie

1 INTRODUCTION

About four years ago, the KCE published a first HTA report on Positron Emission Tomography (PET) scanning [1]. It assessed the clinical indications of PET, the cost-effectiveness, the number of PET scanners needed in Belgium and the financing of PET scan. With 1.3 approved PET scanners per million inhabitants, Belgium still is one of the countries with the highest number of PET scanners [2]. Moreover, many hospitals in Belgium have a non-approved PET scanner, although the exact number is unknown.

The programming of PET scan is regulated through the hospital law of August 7th 1987, the law of April 27th 2005 and some Royal Decrees. However, recently the European Commission received a complaint against these laws, since the programming of PET scan is not based on objective criteria. On January 31st 2008, the European Commission decided to disregard the complaint provided that Belgium adapts its current programming using objective criteria.

In December 2008, the Minister of Health launched an urgent demand to the KCE to update the previous report on PET scan in order to provide a basis for a new programming policy. In her demand, the Minister stated that this new policy should be based on the evolution of the number of patients requiring a PET scan and on evidence-based clinical practice recommendations.

The main objective of the present report is to answer the following research questions (chapter 4 – 7): what is the diagnostic accuracy and clinical effectiveness of PET and PET/CT? What are the clinical indications for PET and PET/CT? Importantly, it is not the intention to develop clinical practice guidelines on PET scan. In chapter 8, an answer is given on the following questions: which programming criteria are used in other countries? Can the number of patients requiring a PET scan be estimated in Belgium? Finally, conclusions and recommendations are formulated in chapter 9.

2 TECHNOLOGY DESCRIPTION

2.1 PET

PET imaging is a non-invasive nuclear medicine examination based on the detection of metabolic abnormalities of disease processes through the use of short-lived radiopharmaceuticals. Where classical imaging techniques give information on the structure and localisation of lesions, PET imaging is used, as a complementary tool, to characterise the function, metabolism, biochemical processes and blood flow of organs and when possible, to detect a greater or lesser radiopharmaceuticals' uptake. To reach this goal, a radioactive isotope is combined with a biochemical substance, active in the tissues. This is the case of glucose becoming 18-Fluoro-deoxyglucose (^{18}FDG) when combined with the positron emitting isotope ^{18}F. Glucose is an interesting tracer because it is absorbed in great amount by cancerous or inflammatory cells. Moreover, the development of vascularisation in the cancerous process reinforces this glucose uptake. Once in the organism, ^{18}FDG emits positrons which annihilates with electrons to produce 2 high-energy gamma photons. These photons are detected by the PET camera and then, an image is produced, to be read by the nuclear medicine specialist.

The determination of a positive result depends on the comparison between a specific region and the adjacent "normal" regions. But certain regions of the body are known to be physiologically glucose avid. Therefore, the categorization of a region with augmented uptake is a very difficult process, based on a careful inspection of the region of interest, contrasting the supposed lesion with the adjacent tissue.

With such a process, the experience of the reader is the most important issue. For that reason, there have been various attempts to make the reading objective, at least in a semi-quantitative way. So far, two techniques are used for that purpose: the Lesion–to-Back-Ratio and the Standardized Uptake Value (SUV). The last one is certainly the most common. It is based on the normalisation of attenuation-corrected images for injected dose and body mass. The SUV is the ratio between the tissue concentration of the radiopharmaceutical (in Bq/g) and the injected dose (in Bq) divided by the body mass (in g). The tissue concentration is evaluated on the scanner with a linear grey scale. The difficulty to standardize the reading of PET images explains why sensitivity and specificity may show such variations for the same indication.

PET and conventional nuclear imaging both are diagnostic radionuclide imaging techniques and involve the use of radiopharmaceuticals (pharmaceuticals labelled with a radioactive isotope). These radionuclides can be localized in a variety of physiological or pathological processes using sophisticated imaging systems. The detection of an abnormal lesion with these modalities is based on the differential radionuclide uptake within the lesion and the surrounding tissues. Whether or not a lesion can be detected is related to the degree of radionuclide avidity, size of the lesion and background activity.

Most radioisotopes used in PET are produced in a cyclotron and once incorporated in biological molecules become positron-emission radionuclides allowing imaging of a variety of physiological or pathological process within the human body. Positrons are positively charged electrons emitted from instable nuclei with an excess of protons. These positrons combine with electrons resulting in pairs of positive and negative electrons which rapidly annihilate converting their mass into energy in the form of two gamma rays travelling at 180° from each other. Modern PET imaging systems are designed for the detection of the simultaneous arrival of each pair of gamma rays and hence, collimators are not required. The location of the emission can be computed as lying on the line connecting the 2 rays and combining results from multiple emissions, an image is constituted with localisation of the sources of emissions. A dedicated PET system consists of a ring detector surrounding the patient and collects the pairs of gamma rays emitted.

The coincident arrival of pairs of gamma rays is subsequently recorded and transformed into images. Compared to gamma cameras, PET has a better spatial resolution and is able to identify lesions typically down to the 7- to 8-mm range.

An external positron-emission source mounted on the PET imaging system allows for attenuation measurement and correction (attenuation refers to the loss of photons through scatter or absorption). This transmission scan is done while the patient remains in position and takes 20 minutes in addition to the time needed for the emission scan. A major limitation of PET is the lack of anatomical details. Therefore, interpretation of PET images requires anatomical information from CT or MRI.

2.2 PET/CT

PET/CT is an emerging technology, where a CT scanner (emitting X-rays) is combined with a PET imager in the same gantry. Typically, the CT acquisition is performed first followed by PET acquisition. The images may then be read separately, or combined using image registration algorithms. This set-up allows co-registration of PET data and CT data producing fusion images with combined functional and anatomical details. In addition, attenuation correction is based on CT data thereby reducing the total scanning time to less than 30 minutes. It has been proposed that PET/CT could be used to improve the PET image through fast and accurate attenuation correction, improve localisation of abnormalities detected on PET, radiotherapy and surgery planning, evaluation of therapy outcome by localising regions of oedema and scarring and produce the highest quality PET and CT information with the least inconvenience. The costs related to the acquisition and the maintenance of a PET/CT scanner may be higher than that of a PET scanner only, but may be outweighed by the potential of producing diagnostically superior images and reducing scan time, thus allowing higher patient throughput.

PET/CT has been reported to be the fastest growing imaging modality worldwide, with standalone PET scanners no longer being produced [3].

2.3 PET TRACERS

Cyclotrons produce the radioisotopes used for PET scanning. The isotopes principally used include oxygen (^{15}O), nitrogen (^{13}N), carbon (^{11}C) and fluorine (^{18}F). Oncological PET tracers are mainly divided into 3 groups: fluorinated tracers (of which FDG is the most frequently used), carbon-11-labelled tracers and other radiotracers.

2.3.1 [^{18}F]fluoro-2-deoxy-D-glucose (FDG)

The most commonly used radiopharmaceutical in PET is an analogue of glucose labelled with ^{18}F (2-deoxy-2-{Fluorine-18}fluoro-D-glucose or FDG) with a half-life of 110 minutes allowing commercial distribution of synthesised FDG within 2 hours. For other isotopes with much shorter half-lives (ranging from 2 minutes for ^{15}O to 20 minutes for ^{11}C), on-site production is required. In this report, for convenience, the term PET is used for FDG-PET unless otherwise specified (e.g. for cardiology, brain tumours and neurology).

The use of FDG is based on the higher rate of glucose uptake in cancer cells caused by an increased expression of transport proteins and upregulation of the hexokinase activity and a decrease in glucose-6-phosphatase activity. After entering the cell, FDG is rapidly phosphorylated to FDG-6-phosphate, which does not cross the cell membrane. Due to its inability to enter the glycolytic pathway and the low levels of glucose-6-phosphatase in cancer cells compared to normal cells, FDG-6-phosphate is preferentially trapped in cancer cells. As this occurs at 50-60 minutes following intravenous administration of FDG, clinical PET imaging is performed after this time interval. In a standard dedicated PET scanner, about 1 hour is required to complete the emission and transmission acquisitions from skull base to thigh. The recent development of faster scintillating crystals and PET/CT systems has reduced total scanning time to less than 30 minutes.

Most frequently clinical PET is used for the detection of lesions and images are qualitatively assessed.

It has been suggested that both attenuation corrected and uncorrected images should be used for lesion detection. While the need for attenuation correction for lesion detection remains debatable, it is certainly required in quantitative measurements of lesion uptake.

However, FDG is non-specific and many inflammatory lesions have also been noted to elevated FDG uptake in PET imaging [4].

2.3.2 Other fluorinated tracers in oncology

^{18}FDG is the most common PET tracer used for the assessment of neoplasms. Many ^{18}F-labelled radiopharmaceuticals other than FDG exploited the characteristics of ^{18}F. They are presented in Table 1, grouped by their mechanism of uptake.

Table 1. Fluorine-18 labelled radiopharmaceuticals and their potential indications

Mechanism of uptake	Radio pharmaceutical	Potential indications
Catecholamine uptake and storage	[^{18}F]fluorodopamine	Neuroectoderm tumours management
Amino acid uptake, decarboxylation and storage	[^{18}F]dihydroxyphenylalanine	Neuroectoderm tumours management
Sympathomimetic amine uptake and storage	[^{18}F]hydroxyephedrine	Neuroectoderm tumours management
Somatostatine receptors mediated	[^{18}F]fluoropropionyl-Lys0-Tyr3-octreotate	Neuroectoderm tumours management
Fluoride ions exchange with hydroxyapatite crystals forming fluoroapatite	[^{18}F]fluoride bone scan	Bone metastases
Biosynthesis of cell membrane component phosphatidylcholine	[^{18}F]choline	Brain, prostate cancer
Diffusion into hypoxic cells	- [^{18}F]fluoroazomycin-arabinofuranoside, - [^{18}F]fluoromisonidazole, - [^{18}F]2-(2-nitro-1[H]-imidazol-1-yl)-N-(2,2,3,3,3-penta-fluoropropyl)-acetamide, - 2-(2-nitroimidazol-1[H]-yl)-N-(3-[^{18}F]fluoropropyl)acetamide	- Tumour hypoxia - Brain, prostate cancer - HNSCC, NSCLC - HNSCC, NSCLC
Estrogen receptors binding	[^{18}F]16 □-fluoroestradiol, [^{18}F]fluorotamoxifen, [^{18}F]fluoro-17-□-estradiol, [^{18}F]fluoro-(2R*,3S*)-2,3- bis(4-hydroxyphenyl) pentanenitrile	Breast cancer
Fatty acid synthesis	[^{18}F]acetate	Prostate cancer
Androgen receptors	[^{18}F]fluoro-dihydrotestosterone	Prostate cancer
Phospholipid synthesis	[^{18}F]fluoroethylecholine, [^{18}F]fluoromethyldimethyl-2-hydroxyethylammonium	Prostate cancer
Thymidylate synthase inhibitor	[^{18}F]5-FU	Colon cancer
Protein synthesis	[^{18}F]fluorotyrosine	Brain tumours
Amino acid transport	[^{18}F]methyl tyrosine	Brain, colon, breast cancer
Transport into cells by thymidine kinase activity	[^{18}F]thymidine	Brain tumours
Binds to externalized phosphatidylserine on apoptotic cells	[^{18}F]annexin V	Various cancers

Adapted from Kumar et al. [4].

2.3.3 Carbon-11-labelled tracers in oncology

The value of ^{18}FDG in the diagnosis of cortical gliomas is limited due to the high physiological uptake in normal grey matter. Therefore, other more specific metabolic tracers with only limited uptake in normal brain tissue, such as positron emitter-labelled amino acids, have been proposed as new predictors. For example, methionine (MET) is a natural essential amino acid and enters tumour cells via the L-amino acid transporter to meet the demands of accelerated protein and RNA synthesis in malignant tumours.

More generally, ^{11}C-labelled tracers have shown high specificity in tumour detection, tumour delineation and differentiation of benign from malignant lesions. Numerous ^{11}C-labelled tracers are presented in Table 2.

Table 2. Carbon-11 labelled radiopharmaceuticals and their potential indications

Mechanism of uptake	Radio pharmaceutical	Potential indications
Glucocorticoid synthesis	[^{11}C]etomidate, [^{11}C]metomidate	Adrenocortical tumours
Catecholamine uptake and storage	[^{11}C]epinephrine, [^{11}C]phenylephrine	Neuroectoderm tumours management
Decarboxylation and formation of biogenic amines dopamine and serotonin	[^{11}C]5-ydroxytryptophan	Serotonin-producing tumours
Neutral amino acid uptake, decarboxylation and storage	[^{11}C]dihydroxyphenylalanine	Neuroectoderm tumours management
Phospholipid synthesis	[^{11}C]choline	Genitourinary cancer and brain tumours
Sympathomimetic amine uptake and storage	[^{11}C]hydroxyephedrine	Neuroectoderm tumours management
Amino acid transport	- [^{11}C]methionine - [^{11}C]tyrosine	- Genitourinary cancer and brain tumours - Brain, colon, breast cancer
Fatty acid synthesis	[^{11}C]acetate	Genitourinary cancer and brain tumours
Nucleoside metabolism and trapping by thymidine kinase	[^{11}C]thymidine	Brain, HNC and lymphoma

Adapted from Kumar et al. [4].

2.3.4 Other radiotracers used in oncology

The other non-FDG radiotracers can be labelled with ^{68}Ga, ^{60}Cu, ^{64}Cu, etc. and are aimed to detect cell hypoxia, bone metabolism and receptor. Many of these have shown promising results in the management of cancers for which FDG had limited value [4]. They are listed in Table 3.

Table 3. Other radiopharmaceuticals and their potential indications

Mechanism of uptake	Radio pharmaceutical	Potential indications
Amino acid transport	[^{124}I]IMT	Brain tumours
Hypoxia	[^{60}Cu]pyruvaldehyde-bis(N4-methylthiosemicarbazone)	HNC, soft tissue sarcoma and uterine cervix cancer
SS receptors mediated	[^{64}Cu]TETA-OC, [^{68}Ga]DOTA-TOC, [^{68}Ga]DOTA-NOC, [^{68}Ga]DOTA-TATE	Neuroectoderm tumours management

Adapted from Kumar et al. [4].

2.3.5 Other tracers used in neurology

2.3.5.1 Parkinson's disease

For the differential diagnosis of idiopathic Parkinson's disease (PD), three-dimensional PET can be a tool to differentiate between normal and abnormal nigrostriatal innervation [5]. PET provides a measure of the in vivo binding and metabolism of compounds that have been tagged with short lived positron emitting isotopes, such as ^{11}C or ^{18}F. Besides diagnostic evaluation in individual patients with unclear PD *diagnosis* or *prognosis*, PET scan has also been used in studies evaluating *neuroprotection* by drugs and detection of *pre-clinical PD*. However, none of these applications have come yet to a stage where its use in routine clinical management of individual patients has been studied.

Currently, metabolic brain imaging with ^{18}FDG and PET has been described as potentially useful in differentiating idiopathic PD from atypical forms.

Another type of radiotracers are brain receptor binding ligands. They are proposed for the same purpose and they include two distinct categories: the presynaptic ligands such as ^{18}F-Dopa, ^{11}C-dihydrotetrabenazine, ^{11}C-CFT, ^{18}F-CFT etc.; and the post-synaptic or D2 receptor radioligands such as ^{11}C-raclopride.

The presynaptic ligands theoretically have the potential to discriminate between PD and other neurological disorders such as essential tremor or Alzheimer's disease. As to the presynaptic tracers, dopaminergic neurons offer three sites to which biological compounds tagged with positron emitting isotopes can bind: the dopamine transporter (e.g. ^{11}C-nomifensine, ^{11}C-CFT, ^{18}F-CFT and ^{11}C-RTI-32 PET), the vesicular monoamine transporter 2 (e.g. ^{11}C-dihydrotetrabenazine PET) and the enzyme aromatic-amino-acid decarboxylase, which is mainly inside the synaptic terminal and enables transformation of dopa to dopamine [5]. The uptake of the radiotracer ^{18}F -dopa is dependent upon all of the above mechanisms. ^{18}F-dopa is a marker of the accumulation and metabolism of levodopa (tagged with ^{18}F) in the putamen and caudate nucleus over the time course of the scan, where the rate of accumulation of ^{18}F -dopa in the striatum is dependent upon the integrity of the terminal plexus [5]. Molecular imaging approaches with ^{18}F-dopa and PET labelling for dopadecarboxylase (the enzyme involved in dopamine synthesis) was considered to be the gold standard for evaluating nigral dopaminergic neurons in Parkinson disease before the advent of dopamine transporter (DAT) tracers [6].

DAT regulates the dopamine concentration in the synaptic cleft through reuptake of dopamine into presynaptic neurons, and can be considered to be a presynaptic ligand. Pharmacologically, DAT serves as the binding site for drugs of abuse (e.g. cocaine and amphetamine) and therapeutic agents (e.g. methylphenidate and bupropion). The density of DAT can be used as a marker for dopamine terminal innervation [6]. DAT radiotracers that have reached phase III or IV of clinical applications include [6]:

- ^{11}C-cocaine,
- [^{123}I] β-CIT (2b-carboxymethoxy-3b-[4-iodophenyl] tropane),
- [^{123}I] FE-CIT (ioflupane),
- [^{123}I]/[^{18}F]/[^{11}C]FP-CIT(N-[3-fluoropropyl]-2ss-carbomethoxy-3ss-[4-iodophenyl]nortropane),
- [^{18}F]/[^{11}C] CFT (2beta-carbomethoxy-3betafl uorophenyl-tropane),
- [^{123}I]/[^{11}C] altropane,
- [^{123}I]/[^{11}C] PE2I (N-{3-iodoprop-(2E)-enyl}-2beta-carboxymethoxy-3beta-{4'methylphenyl} nortropane),
- [^{11}C] methylphenidate.

Besides the category of presynaptic radioligands, the postsynaptic ligands are assumed to allow for discrimination between PD and atypical parkinsonian disorders. Especially the postsynaptic ligand ^{11}C -raclopride has a short half-life of 20 minutes, limiting it's applicability for routine diagnostic purposes. Some new tracers are currently under evaluation, e.g. ^{18}F -desmethoxyfallypride PET (^{18}F -DMFP-PET), a new dopamine D2-receptor ligand with a longer half-life than ^{11}C –raclopride [7].

2.3.5.2 Alzheimer's disease

The use of PET scan in Alzheimer's disease is mainly confined to FDG-PET. Recently, PET scan tests demonstrating AD brain amyloid deposits have been developed, but these promising new tools deserve further diagnostic evaluation.

2.4 ALTERNATIVES TO PET

2.4.1 Gamma Cameras

Gamma cameras are used in conventional diagnostic nuclear imaging procedures in which radionuclides emitting single gamma ray photons are used. Technetium-99m (Tc-99m) is the most commonly used radioisotope that can be added to a variety of pharmaceuticals. These gamma rays are emitted during decay of the radiopharmaceutical and are detected externally by a gamma camera used in a planar or tomographic mode, the latter known as SPECT. The diagnostic information obtained depends on the type and properties of the radiopharmaceutical used. Gamma rays cannot be focused by an optical lens and instead a collimator, a lead plate with an array of small holes, is used to only detect those photons that travel almost perpendicular to the surface of the detector and excluding all other radiation. Therefore, images of the distribution of the radiopharmaceutical obtained with parallel collimators have a low spatial resolution (above 1.5 cm) and lower sensitivity.

Theoretically, dual- or multi-headed planar gamma cameras could be used for PET as an alternative to dedicated PET imaging. However, only few comparative studies with small sample sizes have been performed. Initial studies reported a similar performance of gamma cameras and dedicated PET in the detection of lesions >2 cm but dedicated PET is more accurate in the detection of small lesions.

Gamma camera is not the scope of this project.

3 METHODOLOGY OF LITERATURE REVIEW

3.1 SEARCH QUESTION

The following search question will be addressed in this report: what are the clinical indications for which PET or PET/CT can be used? In order words, the diagnostic accuracy and clinical effectiveness of PET and PET/CT will be assessed. The most appropriate methodology to address this question is that of a systematic review of the literature.

3.2 SEARCH STRATEGY

First, our search focused on HTA reports and systematic reviews published since the previous KCE report [1]. The CRD database (including DARE, the HTA database and NHS EED) was searched in January 2009 using the following search terms in combination: PET:ti, positron:ti and Positron-Emission Tomography (MeSH). In addition, OVID Medline was searched using an adapted version of the Mijnhout strategy in combination with a search filter for systematic reviews and meta-analyses (see appendix). EMBASE was also searched for synthesized evidence (see appendix for search terms). Finally, websites of HTA agencies (see table 4) were searched for additional HTA reports not identified through the above mentioned strategy. The list of consulted websites is a shortened version of that used in the previous KCE report [1], although some additional HTA agencies were consulted (e.g. SBU and IQWIG).

Table 4. Consulted websites of HTA agencies.

HTA agency	Wesbite
SBU	http://www.sbu.se/en/
NICE	http://www.nice.org.uk/
DACEHTA	http://www.sst.dk/english/dacehta.aspx?sc_lang=en
MSAC	http://www.msac.gov.au/
MAS	http://www.health.gov.on.ca/
HAS	http://www.has-sante.fr/portail/jcms/j_5/accueil
AHRQ	http://www.ahrq.gov/
BCBS	http://www.bcbs.com/
CMS	http://www.cms.hhs.gov/
AETSA	http://www.juntadeandalucia.es/salud/orgdep/aetsa/default.asp?V=EN
AATRM	http://www.gencat.cat/salut/depsan/units/aatrm/html/en/Du8/index.html
CCOHTA	http://www.cadth.ca/index.php/en/home
ECRI	https://www.ecri.org/Pages/default.aspx
DIMDI	http://www.dimdi.de/static/de/index.html
IQWIG	http://www.iqwig.de/index.2.en.html

In addition to this search for synthesized evidence, additional primary studies were searched using OVID Medline. Different approaches were used:

- A generic search strategy for oncologic indications was combined with the adapted Mijnhout strategy and specific search filters for diagnostic studies, prognostic studies and randomised controlled trials (RCTs) (see appendix).
- For each tumour type, neurological, cardiovascular and infectious indications, specific search terms (MeSH terms and free text words) were combined with the adapted Mijnhout strategy and specific search filters for diagnostic studies and prognostic studies.
- Finally, for some specific indications (e.g. brain tumours, dementia, Parkinson, cardiology) the FDG-related search terms were removed from the Mijnhout strategy and again combined with specific search filters for diagnostic studies and prognostic studies.

All searches were limited to articles published in English, French or Dutch. A date limit was set between 2005 and 2009. The exact search dates can be found in Appendix 1.

Since it was not the intention to produce clinical practice guidelines on PET and PET/CT, published guidelines were not systematically searched for. Nevertheless, for some tumours guidelines (if available) were used as a reference in the introduction to highlight the current position of PET and PET/CT in the work-up of these tumours. The National Guideline Clearinghouse (www.guidelines.gov) served as a source for these guidelines.

3.3 IN- AND EXCLUSION CRITERIA

Overall, editorials, letters and case reports were excluded.

HTA 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).

A retrospective study design or the presence of differential verification (i.e. more than one reference standard used) were no 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.

For prognostic studies we used the following exclusion criteria:

- Absence of multivariate analysis;
- Use of the index test to modify the management.

The list of excluded studies can be provided on demand.

3.4 QUALITY APPRAISAL

HTA reports and systematic reviews were critically appraised using the INAHTA checklist for the HTA reports and the Dutch Cochrane checklist for the systematic reviews (see appendix). The methodological quality of the diagnostic accuracy studies was assessed with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist (see appendix), which is a standardised instrument endorsed by the Cochrane Collaboration. Finally, RCTs and prognostic studies were critically appraised using the checklists of the Dutch Cochrane Centre (see appendix). All critical appraisals were done by a single KCE expert. However, in case of doubt the quality appraisal was discussed with a second expert.

3.5 DIAGNOSTIC TEST EVALUATION AND LEVELS OF EVIDENCE

Diagnostic tests are used for various purposes: to increase certainty about the presence or absence of a disease; to monitor clinical course; to support clinical management; or to assess prognosis for clinical follow up and informing the patient [8]. Consequently, diagnostic tests have a potential clinical benefit by influencing management, patient outcome and patient well-being. Tests that do not have this potential are obsolete. Moreover, tests that are not sufficiently reliable may cause harm by inducing inappropriate treatment decisions, unnecessary concern or contrarily, unjustified reassurance. The use of diagnostic tests is therefore never neutral and should be considered with proper care.

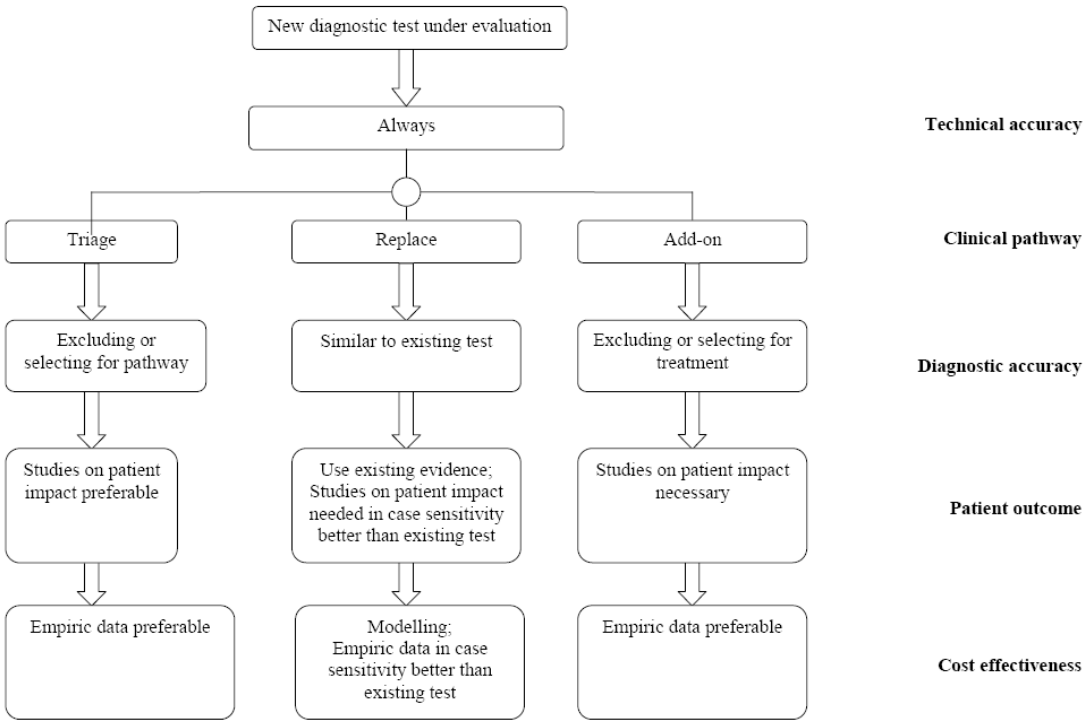
HTA agencies are faced with an increasing demand for the evaluation of diagnostic tests, often after the test has already been introduced in clinical practice. Assessment of diagnostic technologies differs from the evaluation of medical therapeutics in many respects. One of the most important and challenging differences is the indirect relationship between the results of a diagnostic test and the actual health outcome in patients. Diagnostic test results are intermediate outcomes; they influence, but do not directly determine, health outcomes in patients.

The foundation for diagnostic test evaluation was made by Ledley and Lusted in 1959 [9]. Many authors subsequently adopted a hierarchy of diagnostic efficacy with six levels: technical efficacy, diagnostic accuracy, impact on diagnostic thinking, therapeutic impact, impact on patient outcome, and cost effectiveness [10, 11]. But, the intermediate levels either report on information that is already available from a previous level, as is the case with the likelihood ratios and the impact on diagnostic thinking. Or they report on information used as a proxy for the impact on patient outcome, as is the case with the impact on therapeutic management. Another rating scheme has been published by Sackett who identified the four most relevant questions to be asked on a diagnostic test, thereby implicitly ranking evidence [12]. Other authors have stressed the importance of identifying the range of possible uses of the test [13], as this determines what test characteristics the test should have.

In our institution, we have adopted a framework of diagnostic tests evaluation, based on the models proposed by others, but taking the clinical pathway, technical and diagnostic accuracy and patient as well as societal impact into account. The evaluation is stepwise, rather than hierarchical. Every step ought to be taken in order to assess the value of the diagnostic test. The results of previous steps determine the need for evidence in one of the following steps.

Every test evaluation should start with an assessment of the test's capabilities under laboratory conditions. Secondly, the test's place in the clinical pathway should be determined. Thirdly, evidence on the diagnostic accuracy of the test is synthesized according to its intended place in the clinical pathway. Subsequently, the test's impact on the patient is assessed. The final step is a cost-effectiveness analysis, to evaluate the test's value for money as well as other possible societal consequences.

Figure 1: stepwise evaluation of diagnostic tests



Within each step, the evidence can support the use of the test or not, and can be of high or lower quality. Whether a diagnostic test should be implemented in clinical practice, depends on the balance of risks and benefit, and the quality of the evidence underlying this balance, similar to what has been proposed by the GRADE working group [14].

3.5.1.1 Step 1: technical accuracy

The technical accuracy of a test refers to its ability to produce usable information under laboratory conditions and should be done for every diagnostic test under evaluation.

The analytical sensitivity is the ability to detect a specified quantity of the measured component. In these studies, reference material that contains known concentrations of the component of interest is used. Likewise, the analytical specificity is measured on samples that do not contain the component of interest, but contain another component that may cause false positive results.

The reproducibility of results is the ability to obtain the same test result on repeated testing or observations. Reproducibility is influenced by analytical variability and observer interpretation. Analytical variability is due to inaccuracy (systematic error) and imprecision (random error).

Clearly, the test's technical accuracy contributes to its diagnostic accuracy. But there may be a point beyond which improvement in technical performance no longer improves diagnostic accuracy.

3.5.1.2 Step 2: place in the clinical pathway

With the exception of new screening tests, new diagnostic tests fit in an existing pathway. Identification of this existing pathway is crucial in diagnostic test evaluations, as it will determine which characteristics the new test needs to have, but also what information is already available and what is still needed. Recently, three categories were proposed: replacing an existing test in the clinical pathway, before the pathway as a triage, and after the pathway as add-on [13].

A new test may replace an existing test, because it is expected to be more accurate, less invasive for the patient, cheaper, easier to interpret, or yields quicker results. In this situation, the diagnostic accuracy, degree of invasiveness, cost, etc. of the new test will need to be compared with that of the existing test.

Another possible role of a new test could be to triage patients. Typically, triage tests exclude the disease in a proportion of patients, which no longer enter the clinical pathway. However, triage tests may also be used to increase the proportion of patients entering the clinical pathway, by picking up cases that otherwise would have been missed. Triage tests are especially attractive if they are non-invasive for patients, simple to perform and cheap. They need a very high sensitivity to make sure that no cases are missed, because no further tests will be performed on triage-negative people.

Finally, a new test may be placed after an existing clinical pathway, as add-on because it is more accurate or, as the test will only be applied in a subgroup of patients, also more invasive or expensive.

3.5.1.3 Step 3: diagnostic accuracy

Diagnostic accuracy relates to the test's ability to correctly detect or exclude a target condition or disease in patients. Diagnostic accuracy ought to be assessed by synthesizing the available evidence that is appropriate for the intended place in the clinical pathway.

The optimal design is that of the cross-sectional study in which the index test is compared to a reference standard in a cohort of patients that are selected from a clinically relevant population, i.e. patients in whom the test would be applied in clinical practice. This selection is important to avoid spectrum bias affecting estimates of sensitivity and specificity. If sensitivity is determined in seriously ill subjects and specificity in clearly healthy subjects, both will be grossly overestimated [8, 15].

This is an important shortcoming in case-control studies, by which the diagnostic odds ratio is overestimated by a factor 4 in such studies [16].

If a new test is to replace an existing test, head-to-head comparisons of the new and existing test are preferable. The reference standard is performed in all patients and the new and existing test are compared either using a fully paired design, performing the new and the existing test in each patient, or by randomly performing either the new or the existing test. Indirect comparisons might serve as a proxy for paired studies, but should be considered with great caution, as patient population, patient selection and reference standard should be identical in the studies that are being compared.

If the new test will be used to triage patients before the existing pathway, an important aspect is to establish how many diseased cases would not enter the existing clinical pathway. Sensitivity of the triage test can be established by comparing it to the reference standard. But, in order to evaluate in how many patients the existing test can be avoided, it will need to be compared to this existing test as well. Conversely, as the triage test can also lead to an increase in patients entering the clinical pathway, the number of non-diseased entering the clinical pathway and the test's specificity should be assessed as well.

When the test is intended to be used as add-on, the desired test characteristics depend on its goal. Possibly the add-on test should increase sensitivity by decreasing the number of patients testing false negative. Alternatively, the add-on test might be used to increase specificity by decreasing the number of patients testing false positive. The proportion of patients in which treatment is initiated will change, as the add-on test is the final test in the clinical pathway. Provided treatment is initiated in those testing positive, more patients will enter treatment. But, an add-on test may also be used to withhold treatment. In addition, the spectrum of patients who receive treatment will change because a proportion of false negatives and false positives will no longer be treated. Invasiveness, cost, etc. of the add-on test may be compensated by the gain in a better targeted treatment. In conclusion, the add-on test will affect the number and spectrum of patients treated and evidence on patient outcome is necessary.

3.5.1.4 Step 4: impact on patient outcome

The ultimate goal of health care is to improve patient outcome: expected harm, such as burden, pain, risk or costs, should be weighed against expected benefit, such as improved life expectancy, quality of life, avoidance of other test procedures, etc.

The RCT is the study design the least prone to bias to estimate these risks and benefits. However, it is not always feasible to perform an RCT for ethical, financial or other reasons. An important difficulty is that the independent contribution of the diagnostic test to patient outcome may be small in the context of all other influences and therefore very large sample sizes may be required. But, in spite of these difficulties, RCTs on diagnostic tests are feasible. Various designs are possible, depending on the particular research question [17].

If evidence from an RCT is lacking, other study designs may provide some of the answers. One possible design is a controlled trial without randomization; patients who are given the new test are compared to patients who did not. As in all studies using a similar design, attention must be paid to confounding factors and selection bias.

Another study design is a before – after study using a historical control group: data are collected before and after the introduction of a new test. Here, caution is warranted, as other changes might have occurred than merely the change in diagnostic testing. Changes over time in incidence of disease, disease spectrum (e.g. by an advertising campaign to encourage people without symptoms to be tested) or therapeutic advances will also influence patient outcome. Finally, case-control studies can retrospectively give a first idea on the effect of a diagnostic test on patient outcome.

For some tests, however, we will never be able to prove a change in 'objective' patient outcome such as mortality or morbidity, simply because no treatment is yet available that can impact patient outcome, for example in amyotrophic lateral sclerosis (ALS).

A diagnostic test may then improve quality of life by giving the patient an affirmative diagnosis.

When such studies are lacking, studies on the test's influence on the physician's thinking may serve as a proxy. A patient's outcome can not be influenced by diagnostic testing unless the physician is led to do something different than he or she would have done without the test result. Studies can asses the change in diagnosis or intended treatment by the physician, by comparing the intended management before the test result is known to that after the test result has been disclosed. But change in therapeutic management does not necessarily lead to an improved patient outcome. Patients may not benefit from the change in therapy, or even experience harm. In conclusion, studies assessing the test's influence on the physician play only a marginal role in the evaluation process.

Whether new evidence on patient outcome is needed to assess the test's impact, or existing evidence might be used, depends on the test's intended role in the clinical pathway. If a new test is to replace an older, more invasive or expensive test, diagnostic accuracy studies could be a sufficient basis for introducing the test into clinical practice if test characteristics are at least identical, and sufficient evidence on the impact on patient outcome is available for the older test [18]. But, if sensitivity of the new test is better than that of the existing test, new cases are detected and randomized trials are needed as new patients will be entering treatment and the effect on their outcome is unknown [19]. If the test is used as a triage test, RCTs with the new test may not be necessary to evaluate the impact on patient outcome as this remains the same for those entering the existing clinical pathway. But, those patients wrongfully excluded or wrongfully included in the clinical pathway will experience harm, depending on the natural course of the disease in the former case and on the effect of extra testing and treatment in the latter case. In addition, the spectrum of disease of those entering the clinical pathway changes in both cases, leading to changes in the results of the following tests and treatment. As a consequence, new evidence is preferable. When the test will be used as add-on to an existing pathway, RCTs will be necessary, as the spectrum of patients entering treatment will change, or the choice of therapy itself changes depending on the new information.

3.5.1.5 Step 5: cost-effectiveness

Cost-effectiveness analysis goes beyond the individual risks and benefits, but assesses whether the cost of using a given test is acceptable to society.

Cost-effectiveness studies compute a cost per unit of effect measure. Any of a diagnostic test's characteristics can be used as an output parameter, for example cost per surgery avoided, cost per appropriately treated patient, cost per life year gained or cost per quality adjusted life year (QALY) gained. Final outcomes, such as life years gained or QALYs gained, are preferred over intermediate outcomes, as they allow comparisons across a broader range of health interventions, e.g. diagnostic and therapeutic interventions. Because data on these outcomes and costs of the diagnostic and subsequent therapeutic paths are not always available from observations, the cost effectiveness of diagnostic tests is often assessed by means of economic models. The validity of the model's input parameters is crucial for the credibility of the model. The values of all input variables must be based on solid evidence from literature or observations. Sensitivity analyses can demonstrate the sensitivity of the results to changes in the remaining uncertain input parameters. With modelling, it is possible to compute costs per life year gained without having evidence on the impact on patient outcome. However, this approach can be controversial, as many models have to rely on strong assumptions, which are afterwards refuted by observational data.

In economic studies, the use of imprecision estimates, especially the confidence interval, is less established compared to epidemiological and clinical studies. However, economic calculations have their own imprecision that has to be added to the imprecision of the clinical information they are using.

Other societal issues should also be considered when deciding on the proper use of a test in clinical practice, such as equitable access to the test for all patients, consequences for staffing and availability of the test within the broader health care structure.

Based on this stepwise evaluation 4 levels of diagnostic accuracy can be distinguished:

- Level 1: Technical accuracy
- Level 2: Diagnostic accuracy
- Level 3: Impact on patient outcome
- Level 4: Cost-effectiveness

For prognostic studies, no levels of evidence are provided.

Results of diagnostic accuracy studies were assessed according to the intended goal of the PET scan, i.e. whether it was used for inclusion (specificity and positive predictive value) or exclusion (sensitivity and negative predictive value) of a target condition, or the balance between both. Subsequently, the results of the appropriate outcome measure were appraised, along with its imprecision by means of the confidence interval. As a rule of thumb, results >90% were considered good, results between 80-90% moderate and <80% low.

3.6 USED DEFINITIONS IN ONCOLOGY

In oncology, the potential indications for PET and PET/CT cover the entire disease process, ranging from diagnosis over staging to follow-up. However, in the available evidence the used definitions of these stages are not always clear or sometimes confusing. In some studies, patients at different stages of the disease process are included.

To limit the confusion in this report, we consistently used the same term for each stage of the disease process. *Primary diagnosis* was considered to be the detection of the primary tumour before any treatment. *Primary staging* was defined as the evaluation of the extent of the disease in patients with a confirmed tumour before any treatment. In the staging process, one can distinguish nodal staging (or N-staging), distant staging (or M-staging) and evaluation of the resectability.

The stage between neoadjuvant treatment and surgical treatment is a source of confusing terms in the literature. We distinguished the following terms:

- *Restaging*: morphological evaluation of the extent of the disease in patients with a confirmed tumour after neoadjuvant treatment but before surgical treatment.
- *Evaluation of residual mass*: evaluation of the mass that is present after treatment in general (after neoadjuvant treatment, surgery, or adjuvant treatment).
- *Evaluation of treatment response*: evaluation of the (morphological and/or metabolic) response to treatment in general.

For the evaluation of recurrent disease (i.e. during follow-up after treatment), we also distinguished *detection* (or diagnosis) of recurrent disease and *staging* (i.e. evaluation of the extent) of recurrent disease, although this distinction is not always clinically relevant. Recurrent disease was considered a re-appearance of the disease after complete disappearance with treatment.

3.7 EXTERNAL EXPERT MEETINGS

According to the KCE procedures, each report needs to be discussed with experts in the field before it is submitted for final validation. In view of the large variety of possible indications for PET and PET/CT, it was decided to organise four separate external expert meetings. The possible indications were grouped as much as possible according to the organ specialty (see table 5). For each meeting, at least two organ specialists and/or medical oncologists, two radiologists and two nuclear medicine specialists were invited. It was tried to keep a balance between university and non-university affiliations. Each expert received the evidence tables, discussion text and conclusions 7-10 days prior to the meeting. The experts were asked to score their agreement with the conclusions on a scale from 1 (completely disagree) to 5 (completely agree), and to return these scores prior to the meeting.

Table 5. Organisation of external expert meetings.

Expert meeting	Indications	Invited experts
May 15th 2009	Oesophageal cancer, gastric cancer, pancreatic cancer, primary liver cancer, colorectal cancer, GIST	Gastroenterologists (2), radiologists (2), nuclear medicine specialists (2)
May 18th 2009 (1)	Lung cancer, head & neck cancer, thyroid cancer, malignant melanoma	Medical oncologists (2), pneumologists (2), radiologists (2), nuclear medicine specialists (2)
May 18th 2009 (2)	Lymphoma, breast cancer, cervical cancer, ovarian cancer, uterine cancer, renal cancer, bladder cancer, prostate cancer, testicular cancer	Urologists (2), gynaecologists (2), radiotherapist (1), haematologist (1), radiologists (2), nuclear medicine specialists (2)
May 19th 2009	Brain cancer, epilepsy, dementia, Parkinson's disease, cardiology, infectiology	Neurologists (3), cardiologist (1), internists (2), nuclear medicine specialists (3)

4 PET FOR CANCER MANAGEMENT

4.1 LUNG CANCER

4.1.1 Introduction

Patients with lung cancer generally present with symptoms and signs of the tumour (e.g. cough, dyspnoea, weight loss, anorexia, chest pain, haemoptysis and hoarseness). These symptoms are characteristic of lung cancer, but many can be indicative of a number of other diseases. It is possible for a tumour to grow quite large before causing any symptoms. In addition, a proportion of patients is diagnosed after their tumour is picked up incidentally on imaging and may not present with any of the classic symptoms of lung cancer. Solitary pulmonary nodules (SPN) are commonly encountered in clinical practice [20]. By definition, the SPN is a single, spherical, well-circumscribed, radiographic opacity that measures 3 cm in diameter [21]. Determining whether a SPN is benign or malignant is challenging and requires a multidisciplinary approach.

In the previous KCE report, evidence on the role of PET scan was found for initial diagnosis of SPN (> 1 cm) and staging, residual mass evaluation after treatment or detection of recurrent non-small cell lung cancer (NSCLC) [1]. For small cell lung cancer (SCLC), evidence was found for staging and restaging and for pleural disease, evidence was found for diagnosis.

Conclusions of the previous KCE report [1]:

- For malignancy diagnosis of a SPN > 1cm, there is evidence of diagnostic efficacy up to diagnostic thinking based on the existence of a pre-test probability and a likelihood ratio, allowing the computation of a post-test probability. In addition, a post-test probability threshold for cost-effectiveness is provided by economic models: evidence is supportive for the use of PET.
- For the initial staging of a Non Small Cell lung Cancer, there is evidence of diagnostic accuracy. In addition, there is evidence that adding PET to CT is cost-effective, although the incremental benefit in terms of life years gained is small.
- For residual and recurrent disease, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.
- For therapy monitoring, there is a lack of evidence for diagnostic efficacy.
- For irradiated volume optimization, there is a lack of evidence for diagnostic efficacy.
- For staging/restaging SCLC, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.
- For pleural disease, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity. For mediastinal disease, there is no evidence.

4.1.2 Diagnosis of malignancy of a solitary pulmonary nodule

Small pulmonary nodules are common incidental findings on chest imaging. Fast identification of malignant nodules is important because they represent a potentially curable form of lung cancer. Management of SPN varies according to its size (<1cm and >1 cm). PET may be indicated in the initial diagnosis of a SPN >1cm when no clear signs of a benign tumour are found on classical imaging procedures.

4.1.2.1 SPN <1 cm

The previous KCE report found limited evidence on the use of PET for the primary diagnosis of SPN <1cm [1]. The NCCHTA 2007 report [22] identified one HTA report that was already included in the previous KCE report. In accordance with this report, NCCHTA stated that PET would not routinely be used for this purpose without biopsy.

One additional primary study was found including 150 patients with SPN. PET had a sensitivity of 51% (95%CI 34-68%) to detect malignant lesions < 2 cm, with a specificity of 52% (95%CI 34-68%) [23]. No additional primary studies met our inclusion criteria.

4.1.2.2 SPN >1 cm

For patients with a SPN >1 cm, the ACCP recommends that clinicians first estimate the pre-test probability of malignancy either by using clinical judgment or by using a validated model [24-26]. One validated model was developed by investigators at the Mayo Clinic and includes six independent predictors of malignancy: older age, current or past smoking, history of extra-thoracic cancer 5 years before nodule detection, nodule diameter, spiculation and upper-lobe location [27, 28].

The NCCHTA 2007 report [22] identified one HTA report that was already included in the previous KCE report. The previous KCE report itself reported a meta-analysis, pooling 32 primary studies with the calculation of SROC curve [1]. Results were compared with those reported by Gould in 2001. For a median pre-test probability of 40%, the post-test probability in case of a positive result is a little less than 75%. In case of a negative PET result, the post-test probability was 2.7% according to Gould et al. and 4.5% according to the KCE [1].

Wahidi et al. performed a high-quality SR comparing the diagnostic accuracy of PET and dynamic CT [26]. However, no primary studies directly comparing both techniques were identified.

For PET, 17 primary studies (n= 790) were found of which 2 were not included in KCE report [26]. Sensitivity of PET ranged between 80 and 100%, while specificity ranged from 40 to 100% with high variation across studies. For dynamic CT, 7 primary studies (n = 948) were found (none were included in the previous KCE report). Sensitivity of dynamic CT ranged between 98 and 100%, while specificity ranged between 54 and 93%, again with high variation across studies. High variation for specificity may be crucial in patients with a high pre-test probability, because negative PET results do not reliably exclude malignancy. Therefore, the ACCP proposed two different strategies [21]. In patients with a low-to-moderate pre-test probability of malignancy (5 to 60%), the ACCP recommends that PET imaging be performed to characterize the nodule. In patients with an SPN that has a high pre-test probability of malignancy (> 60%), the ACCP suggest that PET not be performed to characterize the nodule. However, according to the Belgian experts consulted for this project, PET is performed regardless of the pre-test probability. In case of a negative PET, the patient is followed-up with CT, even with a high pre-test probability.

Cronin et al. performed a comparison between PET and SPECT [29]. For PET, 22 studies (n = 1 069) were included of which three were not included in the previous KCE report. For SPECT, 7 studies (n = 421) were included. For PET, pooled positive LR and pooled negative LR were 5.44 (95%CI 3.56-7.32) and 0.06 (95%CI 0.02-0.09), respectively. For SPECT, pooled positive LR and pooled negative LR were 5.20 (95%CI 4-6.3) and 0.06 (95%CI 0.04-0.08), respectively. However, according to the Belgian experts, SPECT is not used for this indication in Belgium. Pooled positive LR for other diagnostic tests were 3.91 for CT (95%CI 2.42-5.40) and 4.57 for MRI (3.03-6.1), pooled negative LR were 0.10 for CT (0.03-0.16) and 0.08 for MRI (0.03-0.12) [29].

4.1.2.3 Lung tumour suspected of malignancy

We found 5 primary studies focusing on lung tumours suspected of malignancy, but not clearly defined as SPN. Two focused on FLT-PET [30, 31], one on FDG-PET (using a SUV cut-off) [32] and two performed a direct comparison between FDG-PET(/CT) and SPECT/CT [33, 34]. All studies included less than 50 patients.

For FLT-PET, sensitivity was 84% and 90%, while specificity was 100% in both studies [30, 31]. For FDG-PET, sensitivity was 81% and 100%, while specificity was only 85% and 46% [32, 34]. Ferran et al. found a sensitivity and specificity of 100% and 89% respectively for PET/CT [33]. The two comparative studies showed results concordant with those reported by Cronin et al. Compared to SPECT/CT, PET/CT was found to have a better sensitivity (100% vs. 85%, overlapping confidence intervals), but a similar specificity (both 89%) [33]. Wang et al. found a similar sensitivity between PET and SPECT (both 100%), but a lower specificity for PET (46% vs. 69%, overlapping confidence intervals) [34].

4.1.3 Staging of NSCLC

Staging of confirmed lung cancer is performed before and/or after neoadjuvant therapy, and includes mediastinal staging and metastasis detection. Accurate mediastinal staging is crucial to avoid futile thoracotomy in patients with an option of curative treatment [35].

4.1.3.1 FDG-PET

The previous KCE report stated that there is evidence of diagnostic accuracy for the initial staging of NSCLC [1]. The NCCHTA 2007 report [22] identified two SR and 2 additional primary studies. One SR and one primary study were not included in the previous KCE report. For mediastinal staging, the conclusions from this HTA were largely concordant with those from the previous KCE report. For the detection of metastasis, PET was found to have a sensitivity and specificity of over 90% (apart from brain metastases), although these findings were only based on one primary study [22].

4.1.3.2 FDG-PET/CT

The NCCHTA 2007 report [22] identified three primary studies performed with fusion of images obtained by two separate devices. All studies demonstrated better results for PET/CT than for PET alone, particularly for stage I and II disease.

We found 4 additional prospective studies [36-39] performing a head-to-head comparison between FDG-PET/CT and one other imaging modality, 2 studies evaluating PET/CT [40, 41] and 1 study comparing PET and PET/CT [42]. For mediastinal staging, helical dynamic CT [39] or contrast-enhanced CT [38] did not demonstrate any superiority over PET/CT (table 6), with one study showing a significantly better sensitivity for PET/CT (78% vs. 46%). However, overall sensitivity of PET/CT across the identified studies was low (range 47-86%). In one study MRI (STIR turbo SE imaging) showed comparable diagnostic performances as PET/CT [37].

For metastasis detection, MRI showed comparable sensitivity (68% vs. 71%) and specificity (92% vs. 88%) as PET/CT [37].

Table 6. Primary studies on diagnostic value of PET, PET/CT, CT or MRI for the mediastinal staging of NSCLC.

Study ID	N	Design	PET		PET/CT		CT		MRI	
			Se	Sp	Se	Sp	Se	Sp	Se	Sp
Chin 2007 [39]	134	P	-	-	56% (38-72%)	100% (96-100%)	65% (46-80%)	89% (81-94%)	-	-
Ohno 2007 [37]	115	P	-	-	77% (61-87%)	88% (77-94%)	-	-	91% (77-97%)	93% (84-97%)
Quaia 2008 [38]	150	P	-	-	78% (63-88%)	80% (51-94%)	46% (32-60%)	93% (66-99%)	-	-
Kim 2006* [40]	150	P	-	-	47% (30-65%)	100% (97-100%)	-	-	-	-
Kim 2007 [41]	674	P	-	-	61% (54-68%)	96% (94-97%)	-	-	-	-
Lee 2007 [42]	336[$]	R	61% (43-77%)	94% (90-97%)	86% (67-96%)	81% (71-88%)	-	-	-	-

* All histopathologic subtypes analysed together; [$] 210 patients had PET, 116 other patients had PET/CT.

P = prospective, R = retrospective; Se = sensitivity, Sp = specificity. 95%CI are provided between brackets.

4.1.4 Prognostic value in NSCLC patients

4.1.4.1 Prognostic value at diagnosis

Our search identified one moderate-quality systematic review including 13 primary studies (n =1 474) for the evaluation of the prognostic value of SUV for overall survival in patients with stage I to III/IV disease [43]. The combined HR was 2.07 (95%CI 1.66-2.58) for fixed effects and 2.13 (95%CI 1.54-2.95) for random effects.

Our search identified 6 additional prognostic studies (4 prospective and 2 retrospective) evaluating the prognostic value of PET in patients at various stages of NSCLC (table 7). Three of these studies found PET to have a prognostic value, while three others didn't.

Table 7. Primary studies on prognostic value at diagnosis of PET in patients with NSCLC (using SUV value).

Study ID	N	Outcome	Population	Results
Downey 2007 [44]	487	OS	Patients with R0 resection	No independent predictor of survival (p=0.09) after adjusting for pathologic TNM stage.
Hoang 2008 [45]	214	OS	Advanced stages	No independent predictor of survival (p=0.09) after adjusting for sex, stage and treatment.
Tanvet-Yanon 2008 [46]	59	OS DFS	Advanced stages	PET response (semiquantitative reading) is not prognostic factor of survival (p=0.38).
Shin 2008 [47]	184	OS DFS	Tumour stage patients (N disease detection)	SUVmax is independent prognostic factor of DFS (HR for SUVmax >5: 3.653; p=0.011).
Nguyen 2006 [48]	53	DFS	Stages II and III	SUVmax is an independent predictor of recurrence (p=0.002) and death (p=0.041).
Goodgame 2008 [49]	136	OS R	Stages T1 and T2	High SUV is independently associated with recurrence (p=0.002) and death (p=0.041).

OS = overall survival, DFS = disease-free survival, R = Recurrence.

4.1.4.2 Prognostic value after treatment

Our search identified 3 prospective prognostic studies evaluating the prognostic value of PET in patients after primary treatment (chemotherapy, chemo/RT, RT alone or surgery). Mac Manus et al. found a better 2-year survival (after adjusting for pre-treatment status, weight loss and PET stage) in patients with a complete metabolic response (HR 2.71, 95%CI 1.58-4.7, p=0.0001) [50]. Hoekstra et al. found that metastatic lymph node status measured by PET is a better predictor of survival (HR 2.33; 95%CI 1.04-5.22; p=0.04) than CT (HR 1.87) or FDG-PET metabolic rate of glucose (HR 1.95) [51]. On the other hand, Ohtsuka et al. found that SUV (cut-off value 3.3) did not achieve statistical significance (HR 4.2; 95%CI 0.8-21.5; p=0.079) [52].

4.1.4.3 Prognostic value in recurrent disease

Our search identified one prognostic study evaluating the prognostic value of PET in 62 consecutive patients with suspected recurrence after surgical therapy [53]. SUV was found to be an independent prognostic factor of survival in patients with recurrent disease. In patients with SUV <11 after surgery, median survival was 46 months compared to 3 months in patients with SUV ≥11 (p<0.001).

4.1.5 Monitoring of treatment response in NSCLC

The NCCHTA 2007 report [22] identified six small primary studies (sample size ranging from 25 to 57). For three of these studies, no 2x2 tables could be calculated. Results were found to be discordant.

In addition to the NCCHTA 2007 report, one new primary study of PET scan was identified [54]. However, the study suffered from partial verification.

One additional study (provided by the consulted experts) including 30 patients with stage IIIA-N2 compared PET/CT and remediastinoscopy in order to assess the operability after induction therapy [55]. PET/CT showed a better sensitivity (77%, 95%CI 50%-92%) than remediastinoscopy (29%, 95%CI 41%-68%), although the confidence intervals were overlapping. Specificity for PET/CT was 92%, while specificity for remediastinoscopy was 100%.

4.1.6 Radiotherapy planning in NSCLC

The NCCHTA 2007 report [22] identified four studies already included in the previous KCE report and four additional small primary studies (sample size ranging from 21 to 44). All studies showed that FDG-PET affects radiation volume and dose, but only one study presented results on outcome during follow-up.

No additional primary studies were identified by our search. However, according to the consulted experts, controlled trials are ongoing to evaluate the effect on outcomes.

4.1.7 Detection of recurrent disease in NSCLC

The NCCHTA 2007 report [22] identified one primary study including 42 patients and comparing the diagnostic performance of PET and PET/CT for the detection of recurrent disease. PET/CT showed a better specificity (82%, 95%CI 59-94%) than PET (53%, 95%CI 31-74%) and contributed to a change in management in 12 patients.

4.1.8 Small Cell lung Cancer

The previous KCE recommendation (concerning the use of PET alone) was based on one high-quality HTA report and 4 primary studies [1]. No new eligible studies were identified for PET alone by our present search.

For PET/CT, one recent high-quality HTA report was identified [56]. Seven primary studies were included (two of which were already included in the previous KCE report) evaluating the diagnostic accuracy of PET/CT in the staging and restaging SCLC. Mixed reference standards were used. Sensitivity ranged from 14-100%, while specificity ranged from 78-100% [56].

No additional primary studies were found by our search.

4.1.9 Mesothelioma

Our search identified one prospective study evaluating the prognostic value of FDG-PET in 137 patients with pathologically proven mesothelioma [57]. Flores et al. compared three prognostic factors: high SUV, mixed histology and stage. It was shown that high SUV tumours were associated with a 1.9 times greater risk of death than low SUV tumours (p=0.01). Mixed histology carried a 2.9 times greater risk of death than epithelioid histology (p=0.01). Stages III and IV had a 1.8 times greater risk of death than stages I and II (p= 0.05).

Key messages

- **For the detection of malignancy in patients with a solitary pulmonary nodule, the results of the newly identified evidence are in line with the previous report, where evidence of diagnostic efficacy was found (level 2).**
- **For primary staging, the results for PET are in line with the previous report, where evidence of diagnostic efficacy and cost-effectiveness were found (level 4). No direct comparisons were found between PET and PET/CT, although PET/CT appears to have a similar diagnostic efficacy.**
- **From a moderate-quality systematic review, SUV of the primary tumour appears to be a prognostic factor for overall survival, which is confirmed by most additional primary studies. However, the clinical consequences are unclear.**
- **evidence on the use of PET for the evalution of treatment response or detection of residual tumours is limited, and does therefore not allow the formulation of firm conclusions without taking into account the particular clinical situation (level 2).**
- **Studies show that PET affects radiation volume and dose (level 2). Controlled trials are ongoing to evaluate the effect on outcomes.**
- **The new evidence on the use of PET for the evaluation of recurrent lung cancer is limited to one small study (level 2). Overall, the evidence remains inconclusive.**
- **No eligible studies were identified on the use of PET for the staging or restaging of patients with small-cell lung cancer (level 2)**
- **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).**

4.2 LYMPHOMA

4.2.1 Introduction

Lymphoma is the most common haematological cancer, characterised by malignant changes in lymphocytes. Lymphoma is categorised in Hodgkin's lymphoma and non-Hodgkin's lymphoma. Typically, lymphomas reveal themselves as lumps in lymph nodes, but may also affect other tissues such as skin, liver or gut. In addition, they may produce a variety of other symptoms, such as fever and night sweats [58].

In the previous KCE report on PET scan, evidence on the role of PET scan was found for initial staging and staging of recurrence, residual mass evaluation after treatment, evaluation of treatment response and prediction of relapse and survival. In this update, new evidence on the value of PET scan for the management of patients suspected or diagnosed with lymphoma is summarised.

Conclusions of the previous KCE report [1]:

- PET is not indicated in the initial diagnosis.
- For initial staging and recurrence diagnosis (lymph nodes involvement and extra lymphatic localisation), there is evidence for diagnostic accuracy including the determination of sensitivity and specificity but without mentioning a post-test probability or diagnostic threshold. There are some studies treating changes in patient management but with high heterogeneity.
- For residual mass evaluation, there is clinical evidence up to the diagnostic thinking level because PET allows directing the medical decision on the follow up strategy. There is evidence from one modelling study for cost-effectiveness of PET for re-staging Hodgkin's disease.
- For prognosis, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.

- For evaluation of treatment response, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.

4.2.2 Diagnosis

One HTA report [22] cites a very small study (n=8) on the value of PET for the diagnosis of gastric non-Hodgkin's lymphoma. This study was considered ineligible for the present report (<20 patients). No additional primary studies were identified.

4.2.3 Staging

One systematic review [59], one HTA report [22] and three primary studies[60-62] were identified.

The HTA report [22] found three primary studies that reported sensitivities or specificities, both of which appear to be high. The accuracy of PET and CT was found to be similar in one study. The high sensitivity and specificity of PET for staging is confirmed by one original study additionally identified in our update [62].

The systematic review [59] summarised all evidence on the value of PET for diagnosing bone marrow involvement, and found that the pooled sensitivity was only 51% (95%CI 38-64) and pooled specificity was 91% (95%CI 85-95). Similar results were reported in one primary study [61], i.e. a sensitivity of 65% (95%CI 50-78) and a specificity of 99% (95%CI 95-100). Therefore, the evidence does not support the use of PET for this indication.

Finally, one primary study identified in our update reported the diagnostic accuracy of PET in distinguishing indolent from aggressive types of tumour [60]. The sensitivity of SUV ≥9.5 (not corrected) was 81% (95%CI 61-94%), and the specificity was 81% (95%CI 54-96%). Partial volume corrected SUV ≥11.2 yielded a sensitivity of 81% (95%CI 61-94%) and a specificity of 63% (95%CI 35-85%).

4.2.4 Restaging/monitoring treatment response

One systematic review [63], one HTA report [22] and one primary study were identified [62].

The systematic review [63] found variable results for restaging, both in Hodgkin's lymphoma and in non-Hodgkin's lymphoma. Sensitivities ranged between 60-100%, and specificity between 57-100%. This is confirmed by the original study identified in our update [62], which found a sensitivity of 69% (95%CI 51-83%) and a specificity of 90% (95%CI 80-96%).

In the HTA report [22] studies found a prognostic value of PET in assessing mid-therapy response, although the clinical significance and consequences are not always clear.

4.2.5 Recurrence

No HTA reports or systematic reviews were identified addressing this indication. One primary study was found [62]. In this small study (n=48), sensitivity was 98% (95%CI 87-100%) and specificity was 75% (95%CI 35-97%) for the detection of recurrent lymphoma.

4.2.6 Post-treatment evaluation

One HTA report [22] and two systematic reviews [64, 65] were identified.

Although results are highly variable among studies [64], the accuracy of PET for the assessment of residual disease is low [65]. In patients with Hodgkin's lymphoma, pooled sensitivity was 84% (95%CI 71-92%) and pooled specificity was 90% (95%CI 84-94%). In patients with non-Hodgkin's lymphoma, pooled sensitivity was 72% (95%CI 61-82%) and pooled specificity was 100% (95%CI 97-100%) [65]. The HTA report found similar sensitivity for PET and CT, but higher specificity for PET [22].

4.2.7 Prognosis

One HTA report [22] and three primary studies were identified [66-68]. The study by Gallamini [67] et al. consists of a pooled analysis of two earlier primary studies by Gallamini et al. [69] and Hutchings et al. [70]

In multivariable analyses, PET has an independent prognostic value in predicting progression free survival [67] or failure free survival [68].

One study [66] found low sensitivity (42%) and high specificity (90%) for the prediction of relapse. This is in contrast with the results cited in the HTA report [22].

Key messages

- **No new evidence was found on the diagnostic efficacy of PET for the initial diagnosis of lymphoma.**
- **Based on new evidence, PET has a similar diagnostic efficacy as CT for the staging of lymphoma (level 2).**
- **One new small study showed a low specificity (75%) for the detection of recurrent disease. The evidence is too limited to draw firm conclusions (level 2).**
- **For post-treatment evaluation, results are in line with the previous KCE report (which found evidence of diagnostic accuracy), although heterogeneous (level 2). Intended outcomes are not always specified.**
- **New evidence on the prognostic value of PET is in line with the previous KCE report. However, the clinical consequences are unclear.**

4.3 HEAD AND NECK CANCER

4.3.1 Introduction

Head and neck cancers are a group of related neoplasms that arise in the oral cavity (lip, tongue, gum, floor of mouth, palate), pharynx (oro-, naso- or hypo-pharynx), larynx, nasal cavity and paranasal sinuses, or salivary glands.

Conventional diagnostic and staging procedures include laryngoscopy, oesophagoscopy or endoscopy to identify and evaluate the primary lesion. In cases where the lesion is accessible to biopsy, fine needle aspiration of the primary and any involved lymph nodes is appropriate. CT or MRI may be used to help delineate the extent of the primary tumour and the presence of lymph node metastases. As head and neck cancer includes a number of different malignancies, there is no single staging system [71]. The current options for surveillance after treatment include clinical assessment, CT, MRI and endoscopic examination under anaesthesia with biopsy. Endoscopic assessment with biopsy is still regarded as the gold standard, although this procedure is invasive and may cause morbidity [72].

A high-quality HTA report was published by NCCHTA [22] assessing the clinical effectiveness of PET in head and neck cancer. Management decisions relating to diagnosis, staging/restaging, recurrence, treatment response and RT planning were evaluated separately. Another HTA report was published by MSAC in 2008 assessing the safety, effectiveness and cost-effectiveness of PET and PET/CT for squamous cell cancer of the head and neck [71]: in addition to conventional staging of newly diagnosed or recurrent cancer; in addition to conventional assessment for suspected residual cancer after definitive treatment; and in addition to conventional staging of cancer metastatic to cervical lymph nodes from an unknown primary site. Systematic reviews were also retrieved for staging [73] and detection of recurrences [72, 74, 75] as well as additional primary studies found by our own literature search strategy.

Conclusions of the previous KCE report [1]:

- For diagnosis of an Occult Primary Tumour suspected from a cervical lymph node metastasis when clinical examination, panendoscopy with biopsy and/or conventional imaging modalities (CT/MRI) have failed to identify a primary

tumour, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.

- For diagnosis of an Occult Primary Tumour (suspected from a metastatic carcinoma outside the cervical lymph nodes)
    - suspected from a single metastatic site outside the cervical lymph nodes following an unsuccessful initial diagnostic work up,
    - as well as for the detection or exclusion of additional metastases following an unsuccessful initial diagnostic work up for an Occult Primary Tumour when local or regional therapy is considered as part of a treatment plan for a single metastatic carcinoma outside the cervical lymph nodes,

  there is evidence of diagnostic accuracy.
- For diagnosis of primary head and neck cancer, limited evidence seems supportive for the use of PET in the diagnosis of primary head and neck cancer when CT/MRI results are indeterminate.
- For staging in head and neck cancer, i.e. assessment of regional lymph node involvement, there is evidence of diagnostic efficacy up to diagnostic thinking based on calculated positive and negative likelihood ratios.
- For staging in head and neck cancer, i.e. detection of distant metastases and synchronous primary tumours, there is some evidence of diagnostic accuracy.
- For restaging in head and neck cancer, i.e. assessment of residual or recurrent disease during follow up after treatment, there is evidence of diagnostic efficacy up to diagnostic thinking based on calculated positive and negative likelihood ratios.

4.3.2 Diagnosis

4.3.2.1 Primary diagnosis of head and neck cancer

The NCCHTA 2007 report [22] identified one systematic review that was already included in the previous KCE report [1] and one additional primary study. The additional primary study, conducted in 44 head and neck cancer patients with uncertain clinical evaluation (conventional work-up), found that PET yielded a sensitivity of 92% and a specificity of 65%. Reference standard was neck dissection or biopsy of suspicious areas.

No additional studies were identified by our search.

4.3.2.2 Detection of synchronous primaries

The NCCHTA 2007 report [22] identified the same systematic review as mentioned above and one additional primary study. In this study, 53 patients with newly diagnosed head and neck cancer subjected to clinical exam (endoscopy, CT, X-ray, neck/abdomen US) were included. PET found more second tumours compared with routine methods and discovered distant metastases in two patients, but missed distant metastases in three other patients. No data about TP, FP, TN, FN, Se and Sp were provided.

No additional studies were identified by our search.

4.3.2.3 Detection of occult primary tumour

The NCCHTA 2007 report [22] identified two systematic reviews that were included in the previous KCE report and three small additional primary studies. Two small studies showed that PET had a sensitivity of 63% and 67% and a specificity of 90% and 93% for the detection of occult primary tumours in patients with cervical lymph node metastasis [22]. One other small study compared PET to PET/CT and CT. PET/CT was found to have a better diagnostic efficacy than PET, while PET was found to have a better diagnostic efficacy than CT [22]. However, all 95% confidence intervals were overlapping.

Our search identified one high-quality systematic review on the use of PET and PET/CT for the detection of the primary tumour in patients with carcinoma of unknown primary [76]. Dong et al. searched Medline, Embase and Cancerlit until September 2007. Twenty-one studies on PET were identified.

Pooled sensitivity and specificity were 78% (95%CI 72-84%) and 79% (95%CI 74-83%) respectively. PET detected 29% of the tumours that were not detected with conventional imaging. Eight studies on PET/CT were identified. Pooled sensitivity and specificity were 81% (95%CI 74-87%) and 83% (95%CI 78-87%) respectively. PET/CT detected 31% of the tumours that were not detected with conventional imaging.

In addition to the systematic review of Dong et al., we identified 2 prospective studies [77, 78]. Both studies used histopathology or clinical follow-up as reference standard. One study possibly suffered from incorporation bias [77].

For PET, sensitivity was 86% and 100%, while specificity was 69% and 97%. Both studies reported a change in management in about one fourth of patients [77, 78].

4.3.3 Staging

The MSAC 2008 report [71] included 6 HTA reports published between 2001 and January 2008, of which 2 were published after the previous KCE report [22, 79]. In addition, 3 diagnostic accuracy studies that investigated the additional value of PET or PET/CT in the staging of primary head and neck tumours were included [71]. In the largest study including 134 patients with SCC in the oral cavity, PET was used in addition to CT/MRI. CT/MRI yielded poor sensitivity (31%, 95%CI 17-49%), but good specificity (92%, 95%CI 85-96%) for the staging of lymph node metastasis. The addition of PET increased both sensitivity (57%, 95%CI 39-73%) and specificity (96%, 95%CI 90-99%). In another study, PET/CT was used in addition to conventional imaging (endoscopy, CT, MRI) in 23 head and neck cancer patients prior to tumour resection. For the detection of lymph node metastases, sensitivity remained unchanged at 90% while specificity increased from 75% to 94% (no precision about TP, FP, FN and TN). A third large Taiwanese study in 300 patients with nasopharyngeal cancer found that the sensitivity for the detection of distant metastases increased from 33% (95%CI 21-46%) with conventional work-up alone to 84% (95%CI 72-92%) when PET was added. This increase in sensitivity occurred with no significant decrease in specificity, from 97% (95%CI 94-99%) with conventional imaging to 94% (95%CI 90-96%) with the addition of PET [71].

The MSAC 2008 report also identified a prospective Australian therapeutic impact study [71]. It was found that PET led to a change in management plans in 32% (95%CI 20-46%) of 56 patients. PET detected additional lesions in 36% of patients (20/56). Of those in whom additional lesions were detected, treatment plans changed in 70% (14/20; 95%CI 46-88%), and of those with no additional lesions detected, treatment plans changed in 11% (4/36; 95%CI 3-26%) of patients ($p < 0.001$). Compared to conventional work-up, PET improved pre-treatment staging.

The NCCHTA 2007 report [22] included three systematic reviews that were already included in the previous KCE report and 12 additional primary studies for staging of regional lymph-node involvement. In all these studies the reference standard was histopathology from neck dissection. The results were presented by stage of disease. Four studies showed that PET sensitivity was much lower than that of SLNB in patients with clinically N0 disease. In patients with mixed (T1-T3) or unspecified stages, 8 studies showed that PET or PET + CT had sensitivity of approximately 80% and specificity of 80%-97%. This was comparable to or better than CT or MRI in most studies. One of these studies used SLNB on PET negative necks to improve sensitivity. This combination reduced the number of radical neck dissections from 45 out of 62 compared with 35 out of 62 on CT.

The NCCHTA 2007 report [22] also identified 4 primary studies of PET/CT used in various stages of head and neck cancer. However, none of these provided a patient-based analysis.

Our search identified one additional systematic review [73] and 3 additional primary studies [80-82]. Kyzas et al. performed a meta-analysis of all available studies of the diagnostic performance of PET in patients with HNSCC [73]. They determined sensitivities and specificities across studies and constructed SROC curves using hierarchical regression models. The performance of PET was also compared with that of conventional diagnostic methods (i.e. computed tomography, magnetic resonance imaging, and ultrasound with fine-needle aspiration).

Across 32 studies (1 236 patients with data on lymph node metastases), PET sensitivity was 79% (95%CI 72-85%) and specificity was 86% (95%CI 83-89%). For the subset of cN0 patients, sensitivity of PET was only 50% (95%CI 37-63%), whereas specificity was 87% (95%CI 76-93%). In studies in which both PET and conventional diagnostic tests were performed (24 studies), sensitivity and specificity of PET were 80% (95%CI 72-87%) and 86% (95%CI 82-90%), respectively, and of conventional diagnostic tests were 75% (95%CI 65-83%) and 79% (95%CI 72-85%) respectively. For the subset of cN0 patients, sensitivity of PET was only 52% (95%CI 39-65%), whereas specificity was 93% (95%CI 87-96%). For conventional tests, sensitivity was lower (45%; 95%CI 25-67%) as specificity (87%; 95%CI 72-95%).

Krabbe et al. conducted a prospective study in 38 patients with a newly diagnosed SCC of the oral cavity or oropharynx without signs of cervical lymph node metastasis during physical examination (clinical N0-neck) [80]. PET yielded similar sensitivity (50%, 95%CI 21-78%) than conventional imaging (50%, 95%CI 21-78%), but higher specificity (97%, 95%CI 83-99%; vs. 70%, 95%CI 52-83%). Conventional imaging comprised CT (n=19), MRI (n=10), ultrasonography-guided fine needle aspiration cytology (n=5) or US (n=4).

In a large prospective study including 160 patients with SCC of the oropharynx or hypopharynx, Ng et al. compared the diagnostic efficacy of PET to detect distant metastases with that of multi-detector row computed tomography (MDCT) [81]. Reference standard were either histopathology or clinical exam or imaging follow-up. PET yielded higher sensitivity (77%, 95%CI 56-91%) than MDCT (50%, 95%CI 30-70%) with similar specificity (94%, 95%CI 88-97%; vs. 98%, 95%CI 93-99%). The combination of PET and MDCT increased both sensitivity (81%, 95%CI 61-93%) and specificity (98%, 95%CI 95-99%). Senft et al. also prospectively compared PET and CT for the detection of distant metastases in 92 patients with head and neck squamous cell carcinoma and increased risk for metastases [82]. Reference standard was clinical diagnostic work-up since presentation until a follow-up of 12 months including histopathology.

PET had a higher sensitivity (53%, 95%CI 39-67%) than chest CT (37%, 95%CI 24-52%). Specificities were somewhat similar for PET (93%, 95%CI 86-97%) and for CT (95%, 95%CI 88-98%). The combination of CT and FDG-PET had the highest sensitivity (63%, 95% CI 48-76%), without change in specificity (95%, 95% CI 88-98%). SROC analyses of the five point ordinal scales revealed that the 'area under the curve' (AUC) of PET was significantly higher than that of CT. A same analysis was conducted for the detection of distant metastases and synchronous second primary tumours. In this situation, PET had also a higher sensitivity (58%, 95%CI 45-70%) than chest CT (39%, 95%CI 28-53%). Specificities were somewhat similar for PET (93%, 95%CI 86-97%) and for CT (94%, 95%CI 87-98%). The combination of CT and PET had similarly the highest sensitivity (66%, 95% CI 52-77%), without change in specificity (94%, 95% CI 87-98%).

4.3.4 Restaging/recurrence

The MSAC 2008 report [71] included 6 HTA reports published between 2001 and January 2008, of which 2 were published after the previous KCE report [22, 79]. In addition, 3 diagnostic accuracy studies were identified which reported on the utility of PET for the assessment of suspected residual carcinoma and response to therapy [71]. One small study reported a sensitivity of 83% (95%CI 36-99%) and a specificity of 100% (95%CI 46-100%) for the detection of residual disease with PET/CT in 11 patients where clinical exam and CT showed suspected residual disease. In 23 patients where clinical exam or CT showed suspected residual disease, PET yielded a sensitivity of 83% (95%CI 51-97%) and a specificity of 100% (95%CI 68-100%). Two other retrospective studies of fair quality assessed the diagnostic accuracy of PET in patients with residual disease following definitive treatment [71]. For one of these studies, no 2x2 table could be calculated, while the other study presented an analysis based on the number of hemi-necks instead of a patient-based analysis.

The MSAC 2008 report [71] also identified a therapeutic impact study. PET findings changed management in 21 patients (40%). When the PET scan was negative, the most common change was avoidance of surgery, as reported in 88% of patients (15/17).

The NCCHTA 2007 report [22] included two systematic reviews that were already included in the previous KCE report [1]. In addition, 7 primary studies were included. The seven primary studies included a mixed population of patients with suspected primary or recurrent head and neck cancer. In six studies, PET yielded a sensitivity ranging between 87% and 100% and a specificity ranging between 78% and 97%. These results were better than those obtained for CT (sensitivity 52-67%; specificity 50%) or for CT/MRI (sensitivity 75%; specificity 30%). Only one study showed lower results for PET[22].

For restaging, our search identified one additional primary study [83]. This study compared the diagnostic performance of PET to CT and MRI for restaging after chemoradiotherapy in 31 patients with HNSCC.

Histopathology was considered as reference standard. PET yielded a lower sensitivity (67%, 95%CI 24-94%) than combined CT and MRI (83%, 95%CI 36-99%), whereas its specificity was higher (80%, 95%CI 59-92% vs. 48%, 95%CI 28-68%).

For detection of recurrences, our search identified 3 systematic reviews [72, 74, 84] and one primary study [74].

The systematic review of Brouwer et al. [74] included 8 articles that assessed the diagnostic accuracy of PET in diagnosing recurrent laryngeal carcinoma after radiotherapy. Three studies comprised a comparison with CT and/or MRI (n=181), 1 study was a case-control study. Biopsy taken during direct laryngoscopy and clinical follow-up of 12 months were used as reference standards. In 7 studies, PET yielded a sensitivity of 89% (95%CI 80-94%) and a specificity of 74% (95%CI 64-83%). In one study, PET had a higher sensitivity (80% vs. 58%) but a lower specificity (81% vs. 100%) than CT.

The systematic review by Isles et al. [72] included 27 primary studies that assessed the diagnostic accuracy of PET in diagnosing recurrences in patients with head and neck cancer following radiotherapy or chemoradiotherapy.

Histology from biopsy or surgical specimen and length of disease free survival were used as reference standard. Meta-analysis with random-effects models was conducted. The pooled sensitivity and specificity of PET for detecting residual or recurrent head and neck squamous cell carcinoma were 94% (95%CI 87-97%) and 82% (95%CI 76-86%) respectively. The pooled sensitivity and specificity of PET for detecting residual or recurrent disease of nodal metastasis were 74% (95%CI 50-89%) and 88% (95%CI 74-95%) respectively.

Liu et al. performed a systematic review to compare PET, CT and MRI for diagnosis of local residual or recurrent nasopharyngeal carcinoma [84]. Twenty-one papers were included. The pooled sensitivity estimates for PET (95%, 95%CI 90-97%) were significantly higher than CT (76%, 95%CI 70-81%) ($p<0.001$) and MRI (78%, 95%CI 71-84%) ($p<0.001$). The pooled specificity estimates for PET (90%, 95%CI 87-93%) were also significantly higher than CT (59%, 95%CI 55-63%) ($p<0.001$) and MRI (76%, 95%CI 71-80%) ($p<0.001$). SROC analysis showed better diagnostic accuracy for PET than CT and MRI. PET had significantly better sensitivity and specificity than CT and MRI and was considered as the best modality for diagnosis of local residual or recurrent nasopharyngeal carcinoma.

Brouwer et al. conducted a prospective study including 30 consecutive patients suspected of recurrent laryngeal carcinoma after radiotherapy [74]. All patients underwent PET and direct laryngoscopy under general anaesthesia with taking of biopsies. The suspicion was either raised by symptoms, such as voice deterioration, pain, dyspnoea or dysphagia, or by physical exam (i.e., office laryngoscopy). PET scans were reported by nine blinded nuclear medicine physicians and residents. The reference standard was the absence or appearance of a local recurrence in the 12 months following PET. Sensitivity of PET was 88% (95%CI 53-98%) and specificity was 82% (95%CI 62-93%).

4.3.5 Monitoring of treatment response

The NCCHTA 2007 report [22] found six studies (n=162) that used PET to predict response to therapy (mainly after neoadjuvant therapy). However, since no data about sensitivity and specificity were provided, they were considered ineligible for the present report.

4.3.6 RT planning

The MSAC 2008 report [71] identified the NCCHTA 2007 report [22] and 3 additional therapeutic impact studies. One study reported an overall change in management due to PET/CT in 11 patients (31%). Treatment plan alterations included addition of chemotherapy or radiotherapy so that multimodality approach was favoured. Another study found that in 4/35 patients (11%), PET had a 'high impact' by changing the treatment modality. In 10/35 patients (29%) the impact was considered 'medium' as radiotherapy planning technique or dose was altered.

Finally, in the third study PET was positive in 26/38 cases. In 23 patients, PET/CT provided additional information which altered the treatment plan.

The NCCHTA 2007 report included 3 primary studies (n=47) that used PET in RT planning [22]. PET findings resulted in change in gross tumour volume or the number of irradiated nodes in several patients. The NCCHTA 2007 report also identified three studies (n=88) that used PET/CT for RT planning and showed changes (increase or decrease) in volume or dose compared with CT.

4.3.7 Prognosis

No HTA report or systematic review was found for prognosis. Our search identified only one study [85] investigating 35 successive patients with a clinical or iconographical suggestive SCC of the head and neck (for staging) and followed until death or until last follow-up (prognosis). On univariate Cox regression, only the SUV mean bone marrow activity was predictive of DFS (p=0.05) and OS (p=0.028). When included in a multivariate model with age and gender as covariates, SUV mean bone marrow activity retained its prognostic value for DFS (p=0.04) and for OS (p=0.03).

Key messages

- **New evidence on the use of PET for the detection of head and neck cancer is limited to one primary study, which confirms the previous conclusions (level 2). PET can be used when CT/MRI are equivocal.**
- **A good meta-analysis published since the previous KCE report showed a moderate diagnostic efficacy for PET and PET/CT for the detection of occult primary tumours, although the results were heterogeneous. In view of a detection rate of about 30% in addition to a conventional work-up, PET and PET/CT seem to have a potential role for this indication (level 2).**
- **The diagnostic efficacy of conventional work-up combined with PET is higher than that of conventional work-up alone for nodal and distant staging (level 2).**
- **There is consistent evidence to recommend against a routine use of PET for the nodal staging in patients with head and neck squamous cell carcinoma and a clinically negative neck (level 2).**
- **For the detection of residual disease, new evidence shows moderate for PET, especially for assessing residual N-disease (level 2).**
- **New evidence shows moderate results of PET (pooled sensitivity 89%, pooled specificity 74%) for the detection of recurrence in patients with head and neck cancer (level 2). However, for patients with nasopharyngeal carcinoma, the diagnostic efficacy of PET is significantly higher than that of CT and MRI for the detection of recurrent disease (level 2).**

- **New evidence on the use of PET or PET/CT for RT planning is limited to small studies, by which no firm conclusions can be drawn (level 2).**
- **Evidence on the prognostic value of PET in patients with head and neck cancer is too limited to draw firm conclusions.**

4.4 COLORECTAL CANCER

4.4.1 Introduction

In patients with suspected colorectal cancer, total colonoscopy with biopsy is indicated [86, 87]. For the staging of confirmed colorectal cancer, thoracic and abdominal CT (or MRI) and TRUS (in case of rectal cancer) are recommended. PET(/CT) is recommended for the preoperative evaluation of potentially resectable liver metastases.

Follow-up diagnostic tests after curative treatment include total colonoscopy and lung/liver imaging (with X-ray, CT or MRI). PET can be recommended in patients with a high clinical suspicion of recurrent disease [87].
Conclusions of the previous KCE report [1]:

- For initial diagnosis and staging of colorectal cancer, there is evidence of diagnostic accuracy inlcuding the determination of sensitivity and specificity.
- For restaging after chemo/radiotherapy, there is no evidence.
- For detection and localization of local, hepatic and extrahepatic recurrence, the diagnostic efficacy includes changes in patient management and therapeutic decision. In addition, there is limited evidence for cost-effectiveness.
- For treatment monitoring, there is no evidence.

4.4.2 Diagnosis

The previous KCE report only found limited evidence on the use of PET for the primary diagnosis of colorectal cancer [1].

The NCCHTA 2007 report [22] identified one HTA report that was already included in the previous KCE report. One additional primary study was found including 45 patients with colonic polyps. PET had a sensitivity of 62% to detect malignant lesions, with a specificity of 100% [22].

No additional primary studies met our inclusion criteria.

4.4.3 Primary staging

The previous KCE report identified evidence from 1 HTA report and 2 primary studies for the diagnostic efficacy of PET for the primary staging of colorectal cancer [1].

The NCCHTA 2007 report [22] identified one systematic review that was already included in the previous KCE report. In addition, 2 primary studies were found evaluating the use of PET before initial therapy for colorectal cancer. One of these studies was already included in the previous KCE report, the other study was a therapeutic impact study. A change in management was found in 17% of patients.

The NCCHTA 2007 report also identified 5 primary studies evaluating the use of PET for the staging of patients considered eligible for resection of colorectal liver metastases [22], one of which already was included in the previous KCE report. Most studies included both patients with primary and recurrent disease. Only one study was suitable to calculate sensitivity and specificity. In this study, PET had a better sensitivity than CT (100% vs. 47%) for the detection of liver metastases. Three studies reported a change in management in 9-39% of patients [22].

Finally, the NCCHTA 2007 report identified 2 prospective studies evaluating the use of PET/CT for the staging of colorectal cancer [22]. Only one study reported a patient-based analysis. PET/CT had a similar sensitivity as CT (91% vs. 92%), specificity was better for PET/CT (90% vs. 70%), although the 95% confidence intervals overlapped.

In a recent meta-analysis, Bipat et al. included 21 primary studies (published before December 2003) evaluating the use of PET for the detection of liver metastases [88]. PET was found to be the most accurate imaging modality, with a sensitivity estimate on a per-patient basis of 95%, compared to 60%, 65% and 76%, for non-helical CT (28 studies), helical CT (15 studies) and 1.5-T MRI (12 studies) respectively.

Our search identified 5 additional primary studies. One retrospective study evaluated the use of PET/CT for the N-staging of 53 patients with rectal cancer [89]. For the detection of pararectal nodes, sensitivity was 73% for non-enhanced PET/CT and 90% for contrast-enhanced PET/CT, specificity was 57% and 78% respectively. For the detection of internal iliac nodes, sensitivity was 60% for non-enhanced PET/CT and 73% for contrast-enhanced PET/CT, specificity was 82% and 87% respectively. For the detection of obturator nodes, sensitivity was 50% for non-enhanced PET/CT and 80% for contrast-enhanced PET/CT, specificity was 84% and 91% respectively. All 95% confidence intervals were overlapping [89].

Four primary studies evaluated the use of PET (1 study) and PET/CT (3 studies) for the detection of colorectal liver metastases (see table 8). In the prospective study comparing PET and CT, PET was found to be more sensitive than CT, although the 95% confidence intervals were overlapping [90]. For PET/CT, sensitivity ranged from 94-98%, while specificity ranged from 75-100% [91-93]. Two studies found PET/CT to be as accurate as MRI, although one study suffered from incorporation bias for MRI [91]. In the retrospective study that compared PET/CT to CT, PET/CT was found to be more specific than CT, although the 95% confidence intervals were overlapping [93]. This study also suffered from incorporation bias.

4.4.4 Monitoring of treatment response

The previous KCE report did not identify evidence on this indication [1].

The NCCHTA 2007 report identified 6 primary studies evaluating the use of PET for the monitoring of treatment response in patients with rectal (5 studies) or colorectal cancer (1 study) [22]. Only in 3 studies a 2x2 table with sensitivity and specificity was calculable. Sensitivity and specificity for the differentiation between responders and non-responders ranged from 79-100% and 45-86% respectively [22].

Our search identified 3 additional prospective studies evaluating the use of PET [94] and PET/CT [95, 96] for the evaluation of treatment response in patients with rectal cancer undergoing neoadjuvant chemoradiotherapy (table 9). Depending on the timing of PET(/CT) and the used threshold, sensitivity ranged from 74-100% and specificity from 70-79%.

4.4.5 Radiotherapy planning

No evidence on this indication was identified in the previous KCE report [1].

The NCCHTA 2007 report identified one small primary study evaluating the use of PET for the radiotherapy planning in patients with rectal cancer [22]. A good correlation was found between radiotherapy planning regions produced by PET and CT. No additional primary studies were identified by our search.

Table 8. Primary studies on diagnostic value of PET, PET/CT, CT or MRI for the detection of colorectal liver metastases.

Study ID	N	Design	PET		PET/CT		CT		MRI	
			Se	Sp	Se	Sp	Se	Sp	Se	Sp
Coenegrachts 2009 [91]	24	P	-	-	96% (79-100%)	NC	-	-	100% (86-100%)	NC
Kong 2008 [92]	65	R	-	-	98% (91-100%)	100% (40-100%)	-	-	98% (91-100%)	100% (40-100%)
Chua 2007 [93]	131	R	-	-	94% (85-98%)	75% (35-97%)	91% (82-97%)	25% (3-65%)	-	-
Llamas-Elvira 2007 [90]	104	P	89% (64-98%)	93% (85-97%)	-	-	44% (22-69%)	95% (88-98%)	-	-

P = prospective, R = retrospective; Se = sensitivity, Sp = specificity. 95%CI are provided between brackets; NC = not calculable.

Table 9. Primary studies on diagnostic value of PET for the evaluation of treatment response in rectal cancer.

Study ID	N	Index test		Cut-off ΔSUV	Sensitivity	Specificity
Cascini 2006 [94]	33	PET	Before and during CRT	-42%	100%	87%
Rosenberg 2009 [95]	29	PET/CT	Before and during CRT	-35%	74%	70%
			Before and after CRT	-57.5%	79%	70%
Capirci 2007 [96]	45	PET/CT	Before and after CRT	-66.2%	81%	79%

CRT = chemoradiotherapy; SUV = Standardised Uptake Value.

4.4.6 Detection and staging of recurrent disease

The previous KCE report identified an important body of evidence for the use of PET for the detection and staging of recurrent disease, in particular of colorectal liver metastases [1]. Evidence of changes in patient management and even cost-effectiveness was found.

The 2008 report [3] identified 6 HTA reports, of which the NCCHTA 2007 report (see below) was the most recent. Two additional diagnostic accuracy studies and 1 additional therapeutic impact study were found [3]. In the one diagnostic accuracy study with enough information to calculate the 2x2 table, sensitivity and specificity were 89%. The therapeutic impact study found a change in management in 65% of patients when PET was used for the detection of recurrent disease and 42% of patients when used for the staging of solitary lung or liver metastases [3].

The NCCHTA 2007 report identified 3 systematic reviews that were already included in the previous KCE report [22]. Apart from the 5 primary studies evaluating the use of PET for the staging of patients considered eligible for resection of colorectal liver metastases (using mixed populations with primary and recurrent disease, see 4.4.3), the NCCHTA 2007 report identified 4 additional primary studies that assessed colorectal cancer recurrence. In the 3 studies with available 2x2 tables, sensitivity ranged from 85-100%, while specificity ranged from 83-100% [22].

The NCCHTA 2007 report also identified 3 retrospective studies comparing PET and PET/CT [22]. Two studies reported a patient-based analysis. Sensitivity was 88% for PET in both studies, and 86% and 96% for PET/CT. Specificity was 56% and 74% for PET, and 67% and 89% for PET/CT [22].

Our search also identified 3 systematic reviews [88, 97, 98]. The review of Wiering et al. was of moderate quality, with few details on the statistical methods of their meta-analysis [98]. Above this, the search date was superseded by the two other reviews. Therefore, this study will not be discussed. The study of Bipat et al. is already discussed in the part on primary staging (see 4.4.3). Zhang et al. [97] included 27 primary studies (published before January 2008) evaluating the use of PET for the detection of colorectal cancer recurrence, of which some were also included in the MSAC 2008 report and the NCCHTA 2007 report. For the detection of distant recurrence or whole-body involvement, pooled sensitivity and specificity were 91% (95%CI 88–92%) and 83% (95%CI 79–87%) respectively. For the detection of hepatic recurrence, pooled sensitivity and specificity were 97% (95%CI 95–98%) and 98% (95%CI 97–99%) respectively. For the detection of pelvic or locoregional recurrence, pooled sensitivity and specificity were 94% (95%CI 91–97%) and 94% (95%CI 92–96%) respectively [97].

Our search identified 3 additional primary studies. In a randomised trial, Sobhani et al. randomised 130 patients treated with curative R0 surgery for colon or rectal cancer to follow-up with PET or conventional follow-up without PET [99]. No information was available on the randomisation procedure. PET images were interpreted without knowledge of CT results, but other information on blinding was lacking. An intention-to-treat analysis was used.

For all the patients with a recurrence, the time from baseline until detection of the recurrence was significantly shorter in the PET group than in the conventional group (12.1 ± 3.6 vs. 15.4 ± 4.9 months; p=0.01). Time-to-treatment did not differ significantly (14.8 ± 4.1 months vs. 17.5 ± 6 months; p=0.09). No significant difference in mortality was found.

Votrubova et al. retrospectively compared PET and PET/CT for the detection of recurrent disease in 84 patients that underwent surgery for colorectal cancer [100]. For the detection of recurrence in general, sensitivity and specificity were 80% and 69% respectively for PET, and 89% and 92% respectively for PET/CT. However, 95% confidence intervals were overlapping.

Finally, Bellomi et al. retrospectively compared PET/CT and CT for the detection of recurrent disease in 67 patients treated with radical surgery for rectal cancer [101]. For the detection of local recurrence, sensitivity and specificity were 93% and 98% respectively for PET, and 100% and 98% respectively for CT. For the detection of hepatic recurrence, sensitivity and specificity were 100% for both techniques. All 95% confidence intervals were overlapping.

4.4.7 Prognosis

The MSAC 2008 report [3] identified one ongoing RCT comparing the outcomes of patients with recurrent colorectal cancer selected for surgery on the basis of PET (n=44) vs. patients selected for surgery without PET (n=49). Preliminary results indicated that the 9-month disease-free survival was better in the first group (66% vs. 45%). Nine-month disease-free survival of patients proceeding to hepatic resection did not significantly differ in the PET vs. the non-PET arm (72% vs. 55%, p=0.14). MSAC identified one additional primary study demonstrating that patients with PET-detected disease not apparent on prior imaging have a higher risk of disease progression at 12 months than those without PET-detected extra sites of disease for both colorectal indications (suspected locoregional recurrence, RR 1.67, 95%CI 1.06–2.62; isolated metastases, RR 1.68, 95%CI 1.12–2.52) [3].

Our search identified 4 additional prospective prognostic studies evaluating the prognostic value of PET in patients with colorectal [102] and rectal cancer [96, 103, 104] (table 10). All 4 studies found metabolic response to treatment to be an independent prognostic factor.

Table 10. Primary studies on prognostic value of PET.

Study ID	N	Outcome	Prognostic PET(/CT)-parameter(s)
de Geus-Oei 2008 [102]	61	OS PFS	Metabolic response to palliative chemotherapy
Nakagawa 2008 [103]	59	OS	SUV after neoadjuvant radiotherapy
Capirci 2006 [96]	88	DFS OS	FDG-uptake after neoadjuvant chemoradiotherapy
Kalff 2006 [104]	34	OS TTP	Metabolic response after neoadjuvant chemoradiotherapy

OS = overall survival, PFS = progression-free survival, DFS = disease-free survival, TTP = time-to-progression.

Key messages

- **PET and PET/CT are not indicated for the primary diagnosis of colorectal cancer (level 2).**
- **For the preoperative evaluation of potentially resectable liver metastases in patients with primary colorectal cancer, PET and PET/CT were found to have an moderate to good diagnostic efficacy (level 2).**
- **For the distinction between responders and non-responders to neoadjuvant treatment for rectal cancer with PET(/CT), sensitivity ranged from 74-100%, while specificity ranged from 45-87% (level 2). However, the impact on clinical decision making is yet to be evaluated.**

- **The evidence on the use of PET(/CT) for the radiotherapy planning in patients with colorectal cancer is too limited to draw firm conclusions (level 2).**
- **Based on diagnostic accuracy studies, PET and PET/CT can be recommended for the detection of recurrent colorectal cancer, in particular hepatic recurrence (level 3). One RCT did not demonstrate an effect on mortality of adding PET to a conventional follow-up strategy.**
- **PET can provide prognostic information in patients with colorectal cancer, although the clinical consequences are unknown.**

4.5 MALIGNANT MELANOMA

4.5.1 Introduction

According to the guidelines of the College of Oncology [105], the diagnosis of malignant melanoma is based on a history taking, physical examination, dermatoscopy and diagnostic excision. Preoperative staging is necessary to detect metastases and to have a reference point for post-treatment follow-up. However, no consensus exists on which staging techniques to perform. It is recommended not to use additional staging techniques in case of in situ melanoma. For invasive melanoma, the College of Oncology recommends a chest X-ray, abdominal imaging (US or CT) and imaging of the locoregional lymph nodes (US or CT) [105]. PET(/CT) and brain MRI are considered to be optional.

Conclusion of the previous KCE report [1]:

- For staging in malignant melanoma, i.e. assessment of regional lymph node involvement or distant metastatic disease in patients with primary or suspected recurrent melanoma, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity. Evidence on the use of PET in initial staging is conflicting.

4.5.2 Staging

The NCCHTA 2007 report [22] identified 2 systematic reviews (Mijnhout 2001 and DACEHTA 2001 report) that were already included in the previous KCE report. Both systematic reviews mixed studies with primary and recurrent malignant melanoma. In addition to these 2 reviews, Facey et al. found 12 additional primary studies, mainly on primary malignant melanoma [22]. In patients with early-stage disease (9 studies, n=528), PET was found to have a consistently low sensitivity for regional lymph node detection (range 0-40%). For distant staging, discordant results were found, with 1 study showing a sensitivity of only 4%. In patients with advanced stages, sensitivity ranged from 40 to 100% (3 studies) [22].

In a more recent systematic review, Krug et al. identified 28 primary studies [106]. The pooled sensitivity for the detection of metastasis in patients with early-stage disease was 60%. In patients with advanced stages, pooled sensitivity and specificity were 86% and 87% respectively. Eight studies suggested that PET was associated with 33% disease management changes (range 15–64%).

Our literature search identified 4 additional primary studies. One prospective study confirmed the low sensitivity (14%) of PET/CT for the N-staging of patients with stage I/II cutaneous melanoma [107]. In another retrospective study including 47 patients with advanced stage malignant melanoma, PET/CT was found to have a sensitivity and specificity of 89% and 100% for N-staging [108]. However, this study suffered from incorporation bias. Three studies (1 prospective, 2 retrospective) totalling 210 patients with advanced stages found a sensitivity ranging from 85 to 100% and a specificity ranging from 92 to 100% for the M-staging with PET/CT [108-110]. Two of these studies suffered from incorporation bias.

4.5.3 Detection of recurrence

The NCCHTA 2007 report identified 2 primary studies [22]. One study was a patient management trial that showed a change in management in 34% of patients. The other trial was a diagnostic accuracy study that used a lesion-based analysis. A change in management was found in 30% of patients.

No additional primary studies were identified by our search.

4.5.4 Staging of recurrence

The MSAC 2008 report [111] identified 6 HTA reports (including the NCCHTA 2007 report) and 1 systematic review. In addition to the NCCHTA 2007 report (which did not separate results for primary and recurrent malignant melanoma – see above), 1 patient management trial and 1 diagnostic accuracy study were found. PET and PET/CT were found to have a significantly better sensitivity than CT alone for the N-staging of clinically suspected recurrent cutaneous melanoma (100% vs. 100% vs. 84%). For the M-staging, only PET/CT was found to have a significantly better sensitivity than CT alone (100% vs. 86%). The patient management trial found a change in surgical procedure in 10% of patients and a change from surgery to chemotherapy in 13% of patients.

No additional primary studies were found by our search.

Key messages

- **Evidence consistently shows a low sensitivity for the detection of lymph node metastasis in cN0 melanomas (level 2).**
- **For the detection of distant metastasis in patients with primary and recurrent malignant melanoma, a good balance between sensitivity and specificity was found in advanced stages (level 2).**
- **For the detection of recurrence, the evidence is too limited to draw firm conclusions (level 2).**

4.6 BREAST CANCER

4.6.1 Introduction

Currently, conventional mammography is the technique most widely used for the early detection and localization of breast abnormalities [112]. However, its limited sensitivity and specificity hampered the detection and diagnosis of breast lesions, particularly in patients with dense breast parenchyma and in patients with breast implants or surgical scars. PET scanning, scintimammography (SCM), MRI and US have been proposed for this purpose. Yet, the accuracy of these non invasive diagnostic technologies in excluding breast cancer in women at average risk remains unclear [113].

A high-quality HTA report was published by the NCCHTA [22] assessing the clinical effectiveness of PET in breast cancer. Management decisions relating to diagnosis, staging/restaging, recurrence and treatment response were evaluated separately. Systematic reviews were also retrieved for diagnosis [114], staging [114-116] and monitoring of treatment response [114].

Conclusion of the previous KCE report [1]:

- For diagnosis in patients referred for breast biopsy with abnormal mammogram or palpable breast mass, there is evidence of diagnostic inaccuracy. Benefits do not appear to outweigh risks.
- For staging/restaging in breast cancer, i.e. detection of distant metastatic disease if clinical suspicion for metastatic disease is high at initial diagnosis or when recurrent breast cancer is suspected, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity. Evidence seems supportive for the use of PET.

- For staging in breast cancer, i.e. staging of axillary lymph nodes in patients with no palpable axillary lymph nodes metastases and no evidence of distant metastases, there is evidence of diagnostic inaccuracy. Benefits do not appear to outweigh risks.
- For restaging in breast cancer, i.e. detection of loco-regional recurrence, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity. There is inconclusive evidence that PET is superior to CT/MRI.
- For assessment of treatment response, further diagnostic studies are needed.

4.6.2 Diagnosis

The NCCHTA 2007 report [22] included one systematic review conducted by AHRQ [117] and one small primary study. However, the AHRQ 2001 report included data that were already discussed in the previous KCE report and was updated by the AHRQ in 2006 [113]. Consequently, only the latter systematic review will be summarized here. This systematic review was of high quality, but the quality of all of the included studies was moderate. The objective of the systematic review was to determine if the available non-invasive diagnostic tests (PET, MRI, US, SCM) are sufficiently accurate to exclude malignancy, avoiding women with an abnormal mammogram to undergo biopsy. Ninety-six publications were included: 9 on PET (8 WBS, 1 gamma camera), 45 on SCM, 19 on MRI and 8 on ultrasound. Some publications reported data for more than one test. The reference standard was histopathology obtained after biopsy for all studies. Patients considered were those who had suspicious breast lesions (abnormal mammogram and/or physical examination and/or ultrasound examination). For suspicious lesions, sensitivity of diagnostic tests was higher for MRI (92%) than for US (86%) or PET (82%) [113]. On the other hand, specificity was higher for PET (78%) than for MRI (72%) or US (66%). For non-palpable lesions, only scintimammography was studied, yielding a sensitivity of 68% and a specificity of 85%. To place this information into perspective, the authors reported that an average woman in the United States who has an abnormal mammogram requiring a biopsy for evaluation has approximately a 20% risk of cancer. For women at this average level of risk of cancer after an abnormal mammogram, based upon the tests' negative likelihood ratios, the following statements were made [113]:

- For every 1 000 women who had a negative PET scan, about 924 women would have avoided an unnecessary biopsy, but 76 women would have missed cancers.

- For every 1 000 women who had a negative SCM, about 907 women would have avoided an unnecessary biopsy, but 93 women would have missed cancers. (These numbers are for nonpalpable lesions only; numbers could not be calculated for all lesions.)

- For every 1 000 women who had a negative MRI, about 962 women would have avoided an unnecessary biopsy, but 38 women would have missed cancers.

- For every 1 000 women who had a negative US, about 950 women would have avoided an unnecessary biopsy, but 50 women would have missed cancers.

With these comparative data, the authors concluded that MRI is a more valuable tool than PET to give a diagnosis. However, if a less than 2% risk of having breast cancer with a negative diagnostic test is considered, an acceptable level of risk for a diagnostic test to reliably preclude biopsy, none of these tests was sufficiently accurate to replace biopsy for women at average risk of breast cancer. This interpretation went in the same direction as the conclusions formulated in 2001 [117]. However, the authors recognized that future studies could overturn these findings. For non palpable lesions, data were insufficient to estimate the accuracy of PET, MRI or US. SCM was not sufficiently accurate to avoid biopsy. For palpable lesions, data were insufficient to estimate the accuracy of PET, MRI, US and SCM.

The additional primary study retrieved by the NCCHTA 2007 report [22] compared PET and MRI in 36 women with suspicious lesions on mammography or clinical examination.

In this study, PET yielded lower sensitivity than MRI (76%, 95%CI 52-91% vs. 95%, 95%CI 74-99%) and a similar specificity (73%, 95%CI 45-91%). PET was less accurate to detect smaller lesions (< 10 mm).

The systematic review conducted by Bourguet et al. [114] found no changes since their previous report.

4.6.3 Staging: axillary lymph nodes

The NCCHTA 2007 report [22] included one systematic review that was already included in the previous KCE report [1], and four additional primary studies evaluating PET for staging of axillary lymph nodes. Two studies used ALND with SLNB as reference standard, one study used only ALND and the fourth study used ALND or SLNB plus ALND. When ALND was used as reference, PET yielded a sensitivity that ranged between 40% and 93%, with a specificity that ranged between 87% and 100%. When ALND + SLNB were used as reference standard, sensitivity decreased to 20-50%, while specificity did not change (82-100%). Since prevalence of node-positive disease approximated 33-64%, 36-67% patients with negative PET would have undetected axillary disease if further tests were not undertaken.

The systematic review conducted by Sloka et al. [116] included 19 studies for staging of axillary lymph nodes in patients with diagnosed breast cancer. Due to the high heterogeneity between studies, planned meta-analysis was not performed. Particularly, reference standards were quite different between studies (histology via ALND, SLNB, histology + ALND, SLNB +histology via ALND). In 3 high-quality studies, i.e. studies with broad generalisability to a variety of patients and no significant flaws in research methods, sensitivity ranged between 61% and 94%, while specificity ranged between 80% and 98%. Two of these studies were already included in the previous KCE report.

Four additional primary studies [118-121] were retrieved by our own literature search. Ueda et al. [120] included 183 patients with primary operable breast cancer that underwent PET/CT and AUS followed by SLNB and/or ALND for axillary staging. Using visual assessment of PET/CT images, PET/CT yielded a sensitivity of 58% (95%CI 44-70%) and a specificity of 95% (95%CI 89-98%). When a cut-off of SUV was set at 1.8, sensitivity and specificity were 36% (95%CI 24-49%) and 100% (95%CI 96-100%), respectively. On the other hand, the diagnostic performance of AUS was not so different, with a sensitivity of 54% (95%CI 31-55%) and a specificity of 99% (95%CI 95-100%). By the combination of PET/CT (visual assessment) and AUS to the axilla, sensitivity and specificity evolved to 64% (95%CI 51-76%) and 94% (95%CI 88-97%) respectively.

Veronesi et al. [121] enrolled 236 patients with breast cancer and clinically negative axilla undergoing PET/CT before surgery. In all patients, SLNB was carried out after identification through lymphoscintigraphy. Patients also underwent ALND in cases of positive FDG-PET or positive SNB. The results of PET scan were compared with histopathology of SLNB and ALND. In all, 103 out of the 236 patients (44%) had metastases in axillary nodes. Sensitivity of PET/CT was low (37%, 95%CI 28-47%), but specificity was acceptable (96%, 95%CI 91-99%). Comparatively, sensitivity and specificity of SNB were 96% (95%CI 90-99%) and 100% (95%CI 96-100%), respectively.

Gil-Rendo et al. [118] conducted a prospective study including 275 women with breast cancer. In a first group (150 women), ALND was performed regardless of PET results with the aim of evaluating the sensitivity and specificity of the technique. In a second group (125 women), the axillary examination was complemented by SLNB only in those with no pathological axillary uptake on the PET scan. In the first group of 150 women who had preoperative PET and ALND, the sensitivity and specificity of PET for detecting axillary status were respectively 90% (95%CI 83-97%) and 98% (95%CI 93-99%). PET detected axillary involvement in 64 of 71 patients (7 false negatives) and correctly diagnosed 78 of 79 patients without axillary metastases.

Finally, Kumar et al. [119] conducted a prospective study in 80 women with a histological diagnosis of breast cancer and clinically negative axillary lymph nodes, in order to assess the diagnostic efficacy of PET in detecting axillary lymph nodes. Overall, 36 out of the 80 patients (45%) had metastases in axillary lymph nodes. Sensitivity of PET was very low (44%, 95%CI 28-62%) whereas specificity was good (95%, 95%CI 83-99%).

4.6.4 Staging: metastases

Shie et al. [122] conducted a systematic review including 6 articles comparing PET and bone scintigraphy for the detection of bone metastasis from breast cancer. Three studies presented patient-based data, whereas the 3 other studies reported lesion-based data. Reference standards were CT, MRI or bone biopsy with clinical follow-up longer than 6 months. The pooled patient-based sensitivity and specificity of PET were 81% (95%CI 70-89%) and 93% (95%CI 84-97%), respectively. For bone scan, the pooled sensitivity was 78% (95%CI 67-86%), while specificity was 79% (95%CI 40-95%).

Bourguet et al. [114] found no changes since their previous report.

One additional primary study was identified by our search [123] that compared the diagnostic efficacy of PET and bone scintigraphy for the evaluation of osteoblastic bone metastases in patients with breast cancer. The sensitivity and specificity of bone scintigraphy were 78% (95%CI 64-88%) and 82% (95%CI 65-92%) respectively, and those of PET were 80% (95%CI 66-89%) and 88% (95%CI 71-96%) respectively.

4.6.5 Restaging

The NCCHTA 2007 report [22] included one meta-analysis already included in the previous KCE report [1] and one additional primary study evaluating PET for restaging. This primary study focused specifically on the detection of bone metastases in 15 women with PET and SPECT (8 patients restaging, 7 initial staging). However, only a lesion-based analysis was presented.

4.6.6 Detection of recurrence

The NCCHTA 2007 report [22] included one systematic review and one additional study, both already included in the previous KCE report 2005. No additional study was found by our search.

4.6.7 Monitoring of treatment response

The NCCHTA 2007 report [22] included one systematic review and three additional studies. In the systematic review, 8 studies were included evaluating the value of mid-course PET to predict response to chemotherapy in locally advanced breast cancer. Substantial differences between study protocols hampered the authors to combine results in a meta-analysis. The studies were small (between 5 and 28 patients per study with at least one scan after the start of treatment), and used a variety of monitoring times, a variety of target responses and a variety of monitoring methods.

One primary study included in the NCCHTA 2007 report [22] was performed in 50 women with large or locally advanced primary breast cancer. Reference standard was pathological response from surgery. In all, 8% of the patients had pathological complete response (CR) and 46% had pathological partial response (PR). Ten percent of patients had clinical CR and 52% had clinical PR. For a reduction rate in SUV=79%, sensitivity and specificity were 85% and 83%, respectively. For a reduction rate in SUV=88%, sensitivity and specificity were 100% and 56%, respectively. Two other small primary studies included in the NCCHTA 2007 report [22] showed that mid-cycle PET may also be able to predict clinical response to chemotherapy in metastatic breast cancer.

Two additional studies were retrieved with our own literature search [124, 125]. Schwarz Dose et al. [124] evaluated PET for the prediction of histopathologic response early during primary systemic therapy of large or locally advanced breast cancer. In all, 104 patients with newly diagnosed large (≥3 cm) or locally advanced non inflammatory breast cancer participated in a prospective RCT comparing 2 regimens of preoperative chemotherapy. For the analysis of the diagnostic accuracy of PET, all patients were grouped. According to the various thresholds for relative decrease in FDG uptake (SUV) to predict histopathologic response after the first cycle of chemotherapy (from 20% to 50%; n=69), PET sensitivity ranged from 67% to 93%, whereas specificity ranged from 22% to 70%. After the first cycle, SUV decreased by 51% ± 18% in histopathologic responders, compared with 37% ± 21% in non responders (p<.01).

An additional decrease of 63% ± 19% from baseline was observed after the second cycle in responders, versus 48% ± 19% in non responders (p<.01). After a single cycle of chemotherapy, PET predicted pCR (specimens with no residual invasive tumour) with a sensitivity of 90% and specificity of 74%.

Berriolo et al. [125] investigated 47 women with non-metastatic, non-inflammatory, large or locally advanced breast cancer receiving different regimens of preoperative chemotherapy. PET was evaluated for monitoring response to neo-adjuvant chemotherapy in breast cancer and compared to histopathologic response after completion of chemotherapy. PET was found to have a high sensitivity (91%) and specificity (86%) for the differentiation between pCR and non-pCR patients.

4.6.8 Prognosis

No HTA report or systematic review was found for this topic. Two primary studies were retrieved [126, 127]. Cachin et al. [126] included 47 women with metastatic breast cancer that were treated with a maximum of three cycles of high-dose chemotherapy. PET was used to predict survival. The median follow-up was 87 months and the median survival, 19 months. In a multivariate analysis of predictive factors for overall survival, the PET result was the most powerful and independent predictor of survival: patients with a negative post-treatment PET had a longer median survival than patients with a positive PET (24 months vs. 10 months; p<0.001). The relative risk of death was higher in patients with PET-positive disease (RR 5.3), prior anthracycline treatment (RR 3.3), or with visceral metastasis (RR 2.4).

Emmering et al. [127] conducted a prospective study in 40 patients who were treated with neoadjuvant chemotherapy for locally-advanced breast cancer (LABC). They determined the prognostic value of preoperative PET after neoadjuvant chemotherapy, both as independent indicator and as add-on to postoperative histopathology. Median follow-up was 60 months (range 15-94) and median time-to-progression was 26 months (range 14-90 months). Preoperative PET (HR 4.09; 95%CI 1.26-13.31; p=0.02) was a better indicator for disease-free survival than histopathological examination (HR 2.52; 95%CI 0.77-8.23; p=0.13). In predicting overall survival, both PET (HR 2.77; 95%CI 0.66-11.66; p=0.16) and histopathology (HR 6.53; 95%CI 0.80-53.14; p=0.08) were non-significant predictors. Multivariate Cox regression analysis revealed no added value of histopathology versus PET results.

Key points

- **New evidence confirms the conclusions of the previous KCE report not to use PET for the diagnosis of breast cancer (level 2).**
- **For axillary lymph node staging, new evidence is in line with the previous KCE report, i.e. PET cannot be recommended for this indication (level 2).**
- **Inconclusive evidence was identified on the use of PET for the detection of bone metastases (level 2).**
- **For the detection and staging of recurrence no new evidence was identified since the previous KCE report (level 2). Since PET can inform further management, multidisciplinary discussion is needed in particular cases.**
- **New studies on the evaluation of treatment response show heterogeneous results, by which no firm conclusions can be drawn (level 2).**
- **There is limited evidence that PET can predict disease free survival in locally advanced breast cancer. Above this, there is limited evidence that PET can predict survival in patients with metastatic breast cancer. However, the clinical consequences are unclear.**

4.7 OESOPHAGEAL CANCER

4.7.1 Introduction

In patients with clinically suspected oesophageal cancer, flexible upper gastrointestinal endoscopy with biopsies of suspicious lesions is the diagnostic procedure of choice [128]. When the diagnosis of oesophageal cancer is confirmed, abdominal and chest CT is recommended for the detection of distant metastases and gross invasion of adjacent structures/organs. Endoscopic ultrasonography (EUS) (with fine-needle-aspiration cytology [FNAC] for N-staging) is the diagnostic procedure of choice for locoregional staging in case the CT scan is negative for metastatic disease. In addition, in patients with an option of curative treatment after conventional staging (i.e. CT and EUS), PET(/CT) scan may be considered for the staging of lymph nodes (loco-regional, distal or all lymph nodes) and distant sites other than lymph nodes [128].

Conclusion of the previous KCE report [1]:

- For diagnosis, i.e. the initial detection of a primary tumour, there is lack of evidence.
- For staging in oesophageal cancer, i.e. staging of lymph nodes (loco-regional, distal or all lymph nodes) and distant sites other than lymph nodes, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity. Evidence, although limited, seems supportive for the use of PET.
- For staging in oesophageal cancer, i.e. staging of distant sites, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.
- For assessment of treatment response after patients, eligible for curative surgery, have received neoadjuvant therapy (comparative with initial staging PET result), there is evidence up to diagnostic thinking based on diagnostic accuracy and prognosis.

4.7.2 Diagnosis

Facey et al. [22] identified one HTA report (MSAC 2001) that was already included in the previous KCE report [1]. However, only patients with established oesophageal cancer were included in the 8 primary studies selected by the MSAC report, making a calculation of the specificity of PET impossible.

No additional primary studies were identified by Facey et al. or by our specific search. Therefore, PET is not indicated in the diagnosis of primary oesophageal cancer.

4.7.3 Staging of primary disease

Since the previous KCE report, 2 good HTA reports [22, 129] and 2 moderate-quality systematic reviews [130, 131] were published, including an update of the systematic review of van Westreenen et al. [132]. The latter review will not be discussed, because the search is superseded by the 3 other reviews.

The MSAC 2008 report [129] presented the most recent literature search (up to December 2007) and identified 4 HTA reports (including the NCCHTA 2007 report) and the systematic review of Westerterp et al. [131]. In addition to the NCCHTA 2007 report, 3 new primary studies were identified, all suffering from partial verification.

Apart from 2 systematic reviews included in the previous KCE report (BCBS 2002 report and van Westreenen 2004), Facey et al. identified 4 primary studies [22]. Three of these studies showed low sensitivities for locoregional nodal staging (range 35-55%), but 2 studies showed better sensitivity for M-staging (range 71-88%) [22].

Van Vliet et al. performed a literature search up to January 2006 [130]. Ten primary studies were identified on regional N-staging with PET. Pooled sensitivity and specificity were 57% (95%CI 43-70%) and 85% (95%CI 76-95%) respectively, compared to 80% (95%CI 75-84%) and 70% (95%CI 65-75%) respectively for EUS (31 studies), and 50% (95%CI 41-60%) and 83% (95%CI 77-89%) respectively for CT (17 studies).

The differences between the SROC curves of EUS, CT, and PET for N-staging were not statistically significant. It is not clear from the review how many studies involved a head-to-head comparison.

Van Vliet et al. identified 9 primary studies on distant staging with PET [130]. Pooled sensitivity and specificity were 71% (95%CI 62-79%) and 93% (95%CI 89-97%) respectively, compared to 52% (95%CI 33-71%) and 91% (95%CI 86-96%) respectively for CT (7 studies). The SROC analysis showed that the diagnostic performance of FDG-PET was significantly higher than the diagnostic performance of CT. Again, it is not clear how many studies involved a head-to-head comparison.

Four recent primary studies with different objectives met our inclusion criteria [133-136]. In a retrospective study involving 125 patients with oesophageal cancer without palpable cervical lymph nodes, Schreurs et al. found PET to be equally accurate as ultrasonography for cervical lymph node staging (sensitivity 100% vs. 86%; specificity 97% vs. 100%; overlapping 95%CI) [136]. In 58 patients with superficial oesophageal adenocarcinoma, sensitivity and specificity for PET were 0% and 94% respectively for the detection of lymph nodes [135].

The specificity for the detection of distant metastases was 95% (prevalence 0%). In a small prospective study including 28 patients with cardio-oesophageal cancer, Huguier et al. found a sensitivity of 67% and specificity of 100% for the detection of distant metastases [133]. PET modified the surgical strategy in 2 patients. Finally, Kato et al. found a better diagnostic performance of PET compared to bone scintigraphy for the detection of bone metastases in 44 patients with thoracic oesophageal cancer (sensitivity 92% vs. 77%, specificity 94% vs. 84%), although the confidence intervals were overlapping [134].

Based on this evidence, PET can be considered useful for the initial staging of patients with oesophageal cancer. According to the consulted experts, PET is particularly useful for the staging of patients with a clinical T3-4 tumour, but not for patients with a clinical T1-2 tumour.

4.7.4 Monitoring of treatment response

The NCCHTA 2007 report [22] identified 1 new primary study published since the previous KCE report. PET/CT was found to be more sensitive than CT and EUS for the assessment of complete response after neoadjuvant chemoradiotherapy (87% vs. 27% vs. 20%).

Our search identified 5 additional primary studies. Three studies on PET (2 prospective, 1 unclear; n=106; reference standard histopathology/surgical findings) found a sensitivity of 90-100% and a specificity of 27-90% for the prediction of response to neoadjuvant treatment in patients with locally-advanced oesophageal cancer [137-139]. However, all 3 studies used a different definition of treatment response. Two retrospective studies on PET/CT (n=107; reference standard histopathology) found discordant data [140, 141]. While Roedl et al. found a sensitivity of 91% and a specificity of 90% using a threshold of 63% decrease of PET/CT-volume to predict histopathological response (defined as <10% viable cells in post-surgical tumour specimen) [140], Erasmus et al. found a sensitivity and specificity of only 54% and 45% respectively using visual analysis to detect complete response (defined as no viable cancer in the resection specimen) [141].

During the external expert meeting, several experts mentioned the MUNICON trial as being a relevant study for this indication [142]. This study was indeed identified by our search for primary studies, but excluded due to methodological problems. One-hundred-and-ten patients with locally-advanced adenocarcinoma of the oesophagogastric junction (type 1 or 2) were assessed with PET for metabolic response after 2 weeks of induction chemotherapy. Metabolic responders (defined as having a decrease of SUV of 35% or more) continued to receive neoadjuvant chemotherapy for 12 weeks (3 different possible regimens) and then proceeded to surgery. Metabolic non-responders immediately proceeded to surgery. In this setting, it is impossible to evaluate the diagnostic efficacy of PET to distinguish responders from non-responders.

By providing a different treatment strategy to both groups after PET but before surgery (to be considered as the reference standard), the reference standard does not measure the same response in both treatment groups. Above this, since PET determined this treatment strategy, the study cannot be used as a prognostic study. Indeed, the differences in median overall survival between the metabolic responders and non-responders can be explained by the different preoperative treatment, rather than by the difference in metabolic response after 2 weeks of neoadjuvant chemotherapy.

4.7.5 Detection of recurrent disease

No HTA reports or systematic reviews were identified that addressed or found studies on this indication. Our search identified one primary study (unclear if prospective) including 56 patients with suspected recurrence after definitive treatment (primarily surgery or radiotherapy) of oesophageal squamous cell carcinoma [143]. The reference standard was histopathology or serial imaging (unclear if PET/CT was also involved). Using a patient-based analysis, PET/CT had a sensitivity and specificity of 96% and 55% respectively for the detection of recurrent disease.

4.7.6 Radiotherapy planning

The NCCHTA 2007 report identified 2 small primary studies involving 55 patients in total [22]. In both studies, the use of PET resulted in different radiotherapy target volumes.

4.7.7 Prognosis

The MSAC 2008 report identified 4 primary studies and 1 Australian report on the prognostic value of PET and PET/CT in patients with oesophageal cancer [129]. The evidence for the prognostic value of PET(/CT) was found to be conflicting.

Our search identified 13 additional primary studies, 10 using PET [144-153] and 3 using PET/CT [140, 143, 154] (table 11). All studies used multivariate analysis to detect independent significant prognostic factors. Six studies were prospective (all on PET), 3 were retrospective (2 on PET, 1 on PET/CT), and in 4 studies the study design was unclear (2 on PET, 2 on PET/CT). Only 2 studies found no prognostic value of PET in patients with oesophageal cancer [148, 152]. Complete metabolic response after neoadjuvant treatment was found to be a prognostic factor in 3 studies [149, 150, 153], number of PET-positive lymph nodes in 3 other studies [146, 147, 151].

As mentioned above (see 4.7.4), the MUNICON trial was excluded [142].

Table 11. Primary studies on prognostic value of PET and PET/CT.

Study ID	N	Index test	Design	Outcome	Prognostic PET(/CT)-parameter(s)
Roedl 2009 [154]	49	PET/CT	?	DFS	Decrease of the metabolic tumour diameter between pre- and post-treatment PET/CT
Roedl 2008 [140]	51	PET/CT	R	OS, DFS	Decrease of PET/CT volume
Guo 2007 [143]	56	PET/CT	?	OS	SUV Disease status on PET/CT
Cheze-le Rest 2008 [146]	52	PET	P	OS	SUVmax >9 Number of PET-positive lymph nodes
Chung 2008 [148]	100	PET	?	DFS	-
Makino 2008 [151]	38	PET	R	DFS	Number of PET-positive lymph nodes PET response for primary tumour PET response for lymph nodes
Omloo 2008 [152]	125	PET	P	DFS	-
Kim 2007 [150]	62	PET	P	OS, DFS	Complete metabolic response by PET
Blackstock 2006 [144]	110	PET	P	OS	Metastases detected by PET

Study ID	N	Index test	Design	Outcome	Prognostic PET(/CT)-parameter(s)
Cerfolio 2006 [145]	89	PET	R	OS	SUVmax >6.6
Choi 2006 [147]	51	PET	?	OS, DFS	Number of PET-positive lymph nodes
Duong 2006 [149]	53	PET	P	OS	Complete metabolic response by PET
Ott 2006 [153]	65	PET	P	OS, recurrence	Complete metabolic response by PET

P = prospective, R = retrospective; DFS = disease-free survival, OS = overall survival.

Key messages

- **On the use of PET for the diagnosis of oesophageal cancer no new evidence was found (level 2).**
- **A recent meta-analysis shows that PET has a low sensitivity and moderate specificity for the detection of locoregional lymph nodes, but a moderate sensitivity and good specificity for the detection of distant disease. Therefore, PET can be useful for the initial staging of patients with oesophageal cancer (level 2).**
- **For the evaluation of treatment response, the newly identified evidence does not clearly support the previous recommendations. Definitions of treatment response, study designs and results were found to be heterogeneous (level 2). Since this is a potential indication for PET, good prospective trials and RCTs are needed.**
- **For the detection of recurrent disease, the evidence is too limited to draw firm conclusions (level 2).**
- **PET can provide prognostic information in patients with oesophageal cancer, although the clinical consequences are unknown.**

4.8 STOMACH CANCER

As in oesophageal cancer, in patients with clinically suspected gastric cancer, flexible upper gastrointestinal endoscopy with biopsies of suspicious lesions is the diagnostic procedure of choice [128]. When the diagnosis of gastric cancer is confirmed, abdominal and chest CT is recommended for the detection of distant metastases. EUS with or without FNAC can be considered in patients to be treated with curative intent based on clinical presentation and/or CT. Other imaging techniques, such as PET scan or MRI, are not part of the routine diagnostic and staging work-up of patients with gastric cancer.

The MSAC 2008 report identified no HTA reports or primary studies on PET or PET/CT in patients with gastric cancer [129]. Our search identified 3 eligible primary studies. Sun et al. found a sensitivity and specificity for PET/CT of 86% and 78% respectively for the detection of recurrent disease in 23 patients with surgically treated gastric cancer (with or without suspicion of recurrence) [155]. However, this study had a retrospective design and potentially suffered from incorporation bias.

Two prospective studies evaluated the diagnostic performance of PET for the evaluation of treatment response (1 study after neoadjuvant treatment, 1 study after primary chemotherapy) [156, 157]. Using a cut-off of -35% of SUV decrease between the PET at baseline and 6 weeks after the start of primary chemotherapy, sensitivity and specificity were 83% and 75% respectively for the prediction of complete or partial response [157]. Using the same cut-off of SUV decrease between the PET at baseline and after neoadjuvant treatment, Ott et al. found a sensitivity and specificity of 69% and 82% respectively for the prediction of histopathologic response (<10% residual tumour cells) [156]. In the latter study, multivariate analysis showed metabolic response to be the only significant pre-surgical predictor for survival (p=0.045; RR 0.39, 95%CI 0.16-0.98).

Key messages

- **The evidence on the use of PET(/CT) is too limited to draw firm conclusions (level 2).**

4.9 THYROID CANCER

4.9.1 Introduction

The diagnosis of thyroid cancer is based on a history taking, physical examination and (US-guided) FNAC [158]. For differentiated thyroid cancer, post-treatment follow-up consists of Tg measurement (after TSH stimulation), neck ultrasound and 131I-WBS. At present, PET(/CT) has no clear role in the routine follow-up of patients with differentiated thyroid cancer.

Conclusion of the previous KCE report [1]:

- For restaging, i.e. detection of recurrence of epithelial thyroid cancer in previously treated patients with elevated biomarkers not confirmed by 131I scintigraphy, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.
- For restaging, i.e. detection of recurrence of medullary thyroid cancer in previously treated patients with elevated biomarkers not confirmed by other imaging, there is some evidence of diagnostic accuracy.
- For restaging, i.e. detection of recurrence of thyroid cancer (no differentiation between epithelial and medullary) in previously treated patients without elevated biomarkers and no evidence of disease by 131I scintigraphy but with clinical suspicion of recurrence, there is some evidence of diagnostic accuracy.
- For restaging, i.e. detection of recurrence of thyroid cancer (no differentiation between epithelial and medullary) in patients with otherwise established neoplastic foci, there is some evidence of diagnostic accuracy.

4.9.2 Diagnosis

The NCCHTA 2007 report identified 1 primary study including 43 patients with suspicious thyroid nodules [22]. For the detection of malignant nodules, PET was found to have a sensitivity and specificity of 100% and 63% respectively.

Our search identified 2 additional prospective primary studies involving a total of 86 patients with thyroid nodules and inconclusive cytological results [159, 160]. For the detection of malignant nodules, sensitivity was 100% in both studies. Specificity was found to be only moderate (see table 12). PET was found to avoid 39 to 66% unnecessary thyroidectomies.

Table 12. Primary studies on diagnostic value of PET for the detection of malignant thyroid nodules.

Study ID	N	Index test	Design	Sensitivity	Specificity
Sebastianes 2007 [159]	42	PET	P	100% (95%CI 72-100%)	39% (95%CI 22-58%)
de Geus-Oei 2006 [160]	44	PET	P	100% (95%CI 54-100%)	66% (95%CI 49-80%)
Kresnik 2003 [161]	43	PET	P	100% (95%CI 79-100%)	63% (95%CI 42-81%)

P = prospective, R = retrospective.

4.9.3 Restaging after treatment

The NCCHTA 2007 report identified 4 primary studies totalling 92 patients [22]. Overall, these studies appeared to be of low quality, involving different patient populations and using different reference standards. Therefore, conclusions on the diagnostic value of PET for restaging after treatment are difficult to make.

4.9.4 Detection of recurrence

The previous KCE report already identified evidence of diagnostic accuracy for the detection of recurrent disease in patients with previously treated epithelial thyroid cancer and elevated biomarkers not confirmed by 131I scintigraphy, in patients with previously treated medullary thyroid cancer and elevated biomarkers not confirmed by other imaging, in patients with previously treated thyroid cancer (no differentiation between epithelial and medullary) and clinical suspicion of recurrence but without elevated biomarkers and no evidence of disease by 131I scintigraphy, and in patients with thyroid cancer (no differentiation between epithelial and medullary) and otherwise established neoplastic foci [1].

The NCCHTA 2007 report identified 2 systematic reviews that were already included in the previous KCE report [22]. In addition, 5 primary studies were identified evaluating PET in previously treated patients with well-differentiated thyroid cancer and a suspicion of recurrence [22]. In 3 of these studies, the suspicion was based on elevated Tg levels and negative 131I scintigraphy. Mixed reference standards were used across these 5 studies. PET was found to have a sensitivity of at least 80% and a specificity ranging from 25-83%.

Two additional primary studies evaluated PET in previously treated patients with medullary thyroid cancer and elevated calcitonin or CEA levels [22]. Sensitivity for the detection of recurrent disease varied between the 2 studies (41% and 95%), although PET consistently detected more lesions than other methods.

Finally, the NCCHTA report also identified 2 primary studies evaluating PET/CT in patients with suspected recurrence of differentiated thyroid cancer (elevated thyroglobulin and negative 131I scintigraphy) [22]. Both studies suffered from partial verification (therefore, the results are not presented here).

Our search identified 8 additional primary studies: 4 on PET [162-165], 2 on PET/CT [166, 167] and 2 on both PET and PET/CT [168, 169]. One of the latter 2 studies compared the diagnostic accuracy of FDG-PET, 124I-PET and 124I-PET/CT [166]. However, no patient-based analysis was available for 124I-PET/CT. Of the 8 identified studies, the majority had a retrospective design (see table 13). One study suffered from incorporation bias [169], in 3 other studies possible incorporation bias was present [164, 166, 168]. In 1 study, the reference standard was unclear in 45/108 patients [163]. Two studies compared PET(/CT) to CT [168, 169].

The sensitivity of PET ranged from 62 to 96%, while the specificity ranged from 75 to 100% (table 13). For PET/CT, sensitivity ranged from 68 to 98%, while specificity ranged from 81 to 100%. In the 2 studies that compared PET to CT, no significant differences were found (no overlapping 95% confidence intervals) [168, 169]. In the 1 study that compared PET, PET/CT and CT, PET/CT tended to have a better diagnostic accuracy, although the confidence intervals were again overlapping [169].

Two of these 8 primary studies specifically reported on changes in treatment. Freudenberg et al. found a change in treatment in 25% of patients with PET compared to CT and in 14% of patients with PET/CT compared to CT plus PET [169]. Shammas et al. reported a change in management in 44% of patients with PET/CT [167].

4.9.5 Monitoring of treatment response

The NCCHTA 2007 report identified 1 primary study evaluating the use PET for the monitoring of treatment response [22]. A non-significant trend towards lower FDG-uptake at 3 months was found in tumours with better long-term outcome. However, no cut-off (and sensitivity/specificity) was presented to differentiate responders from non-responders.

No additional primary studies were identified by our search.

4.9.6 Prognosis

No HTA reports or systematic reviews addressed the prognostic value of PET in patients with thyroid cancer. Our search identified one retrospective prognostic study including 400 patients with follicular cell-derived thyroid carcinoma [170]. Multivariate analysis showed that FDG-status (RR 7.69; 95%CI 2.17–24.4) and the number of FDG lesions (RR 1.1; 95%CI 1.08 –1.15) significantly correlated with overall survival.

Key messages

- **PET appears to have a good sensitivity for diagnosing malignancy in thyroid nodules with inconclusive cytological results (level 2). However, further studies are needed to confirm this.**
- **For the detection of recurrent disease, results for PET and PET/CT are too heterogeneous to draw firm conclusions (level 2).**

Table 13. Primary studies on diagnostic value of PET, PET/CT or CT for the detection of recurrent thyroid cancer.

Study ID	N	Design	PET		PET/CT		CT	
			Se	Sp	Se	Sp	Se	Sp
Freudenberg 2008 [168]	21	?	80% (56-94%)	100% (3-100%)	-	-	75% (51-91%)	100% (3-100%)
Freudenberg 2007 [169]	36	R	91% (71-99%)	79% (49-95%)	95% (77-100%)	100% (77-100%)	77% (55-92%)	71% (42-92%)
Alzahrani 2006 [162]	50	R	62% (46-76%)	88% (47-100%)	-	-	-	-
Choi 2006 [163]	108	R	94% (85-98%)	78% (63-89%)	-	-	-	-
Iagaru 2006 [164]	21	R	88% (64-99%)	75% (19-99%)	-	-	-	-
Pryma 2006 [165]	44	R	96% (79-100%)	95% (75-100%)	-	-	-	-
Finkelstein 2008 [166]	65	R	-	-	98% (88-100%)	81% (58-95%)	-	-
Shammas 2007 [167]	61	R	-	-	68% (51-82%)	83% (61-95%)	-	-

P = prospective, R = retrospective; Se = sensitivity, Sp = specificity. 95%CI are provided between brackets.

4.10 PANCREATIC CANCER

4.10.1 Introduction

In patients with a clinical suspicion of pancreatic cancer, diagnostic imaging with abdominal CT is recommended [171]. Patients with a strong suspicion of pancreatic cancer but a negative CT should undergo an EUS. In some cases (e.g. cystic lesions), other imaging techniques (US, PET, MRI, ERCP) can be considered for diagnostic reasons.

When the diagnosis of pancreatic cancer is confirmed and no metastases are found on CT, EUS (with or without FNAC) is recommended to further evaluate the resectability of the tumour [171]. In patients with an option for curative treatment after conventional staging with abdominal CT and EUS, PET(/CT) can be considered for the staging of lymph nodes and distant sites other than lymph nodes.

Conclusion of the previous KCE report [1]:

- For diagnosis, i.e. the detection of pancreatic cancer, there is limited evidence of diagnostic accuracy including the determination of sensitivity and specificity. The clinical utility and advantage over other imaging techniques remain to be established.
- For staging, i.e. detection of metastatic disease, there is limited evidence of diagnostic accuracy including determination of sensitivity and specificity. The clinical utility and advantage over other imaging techniques remain to be established.
- For restaging, i.e. detection of residual or recurrent disease, there is lack of evidence.

4.10.2 Diagnosis

The AHRQ 2008 report identified 7 prospective studies (n=479; reference standard histology/biopsy or clinical follow-up) examining the role of PET in the primary diagnosis and staging of pancreatic cancer [56]. Sensitivity ranged from 73 to 97%, while specificity ranged from 41 to 97%. In 4 other prospective studies totalling 230 patients and with mixed reference standards, sensitivity ranged from 69 to 91% and specificity from 65 to 100% for the primary diagnosis of pancreatic cancer. Finally, in 3 prospective studies on PET/CT totalling 193 patients, sensitivity was 89% in all individual studies, while specificity ranged from 64 to 90%.

Our search identified 3 additional primary studies (2 prospective studies, 1 study with an unclear study design). All studies used histopathology or clinical/imaging follow-up as reference standard, although it is unclear which imaging techniques were included in the follow-up (potential incorporation bias). In 41 patients with a mass-forming pancreatic lesion detected by US, CT or MRI, Singer et al. found a sensitivity of 86% and specificity of 79% for the diagnosis of malignancy by PET [172]. Sperti et al. investigated the diagnostic performance of PET in 50 patients with a suspected cystic pancreatic tumour or intraductal papillary mucinous tumour [173]. PET was found to have a sensitivity and specificity of both 94%. Finally, in 46 patients with a solid pancreatic lesion of at least 1 cm, PET/CT had a sensitivity and specificity of 89% and 74% respectively for the detection of malignancy [174]. There were no significant differences with EUS, ERCP and US, although not all patients received all comparator tests.

4.10.3 Staging

The AHRQ 2008 report identified 8 prospective studies with the combined outcome diagnosis and staging (5 on PET, 1 on PET/CT, 2 on both) [56]. However, the results for diagnosis and staging were not reported separately. In addition, one prospective and one retrospective study were found on staging specifically [56]. In the prospective study, sensitivity and specificity were 81 and 88% respectively for the detection of liver metastases.

Our search identified 3 additional primary studies. In a retrospective study including 50 patients with pancreatic cancer, contrast-enhanced PET/CT had the best specificity for evaluating the resectability compared to PET and non-enhanced PET/CT (although the confidence intervals were slightly overlapping) [175]. This study possibly suffered from incorporation bias. Another small prospective study found a sensitivity and specificity of 70 and 83% respectively for the detection of distant metastases [133]. Again, possible incorporation bias was present. Finally, Nishiyama et al. evaluated the staging accuracy of additional delayed PET imaging (interval of 1h) in 55 patients with confirmed pancreatic cancer [176]. For nodal staging, sensitivity and specificity were 70% and 97% respectively. For detection of distant metastases, sensitivity and specificity were 61% and 100% respectively. This study suffered from incorporation bias.

The consulted experts stressed that PET(/CT) can have a role in the staging of patients with pancreatic cancer in case of a diagnostic conflict after conventional staging. However, these cases need to be discussed first in the multidisciplinary team meeting.

4.10.4 Detection of recurrence

The AHRQ 2008 report identified 1 primary study comparing PET and CT/MRI for the detection of recurrent pancreatic cancer [56]. PET was found to be more sensitive than the combined comparator CT/MRI (96% vs. 39%).

No additional primary studies were identified.

According to the external experts, the detection of recurrent pancreatic disease represents a difficult diagnostic problem. No technique has an acceptable diagnostic efficacy, and in the absence of good alternatives, PET(/CT) should be considered for this indication. Importantly, this is not yet supported by good evidence.

4.10.5 Prognosis

No HTA reports or systematic reviews addressed the prognostic value of PET in patients with pancreatic cancer. One eligible prospective study was identified by our search. In 65 patients with pancreatic cancer, the retention index (100% x [SUV at 2h – SUV at 1h]/SUV at 1h) was an independent prognostic factor of overall survival [177].

Key messages

- **Evidence published since the previous KCE report confirms the conclusions on diagnosis of pancreatic cancer. The clinical utility and advantage over other imaging techniques remain to be established (level 2).**
- **The potential use of PET in case of diagnostic uncertainty (cystic pancreatic lesions, chronic pancreatitis) is confirmed in one additional primary study (level 2).**
- **Evidence on the use of PET for the detection of lymph node and distant metastasis shows a low sensitivity (level 2). One additional study found a high sensitivity for the evaluation of the resectability of pancreatic cancer (level 2). Therefore, until the results of the latter study are confirmed, no firm conclusions can be made on the use of PET for staging of pancreatic cancer,**
- **No additional evidence was found on the use of PET for the detection of residual or recurrent disease.**
- **PET might have prognostic value in patients with pancreatic cancer. However, this remains to be confirmed in good prospective studies.**

4.11 PRIMARY LIVER CANCER

The previous KCE report [1] concluded that in 2005 the clinical evidence did not support the use of PET for diagnosing malignancy of a primary liver tumour.

Two primary studies were identified evaluating the use of PET [178] and PET/CT [179] for the distant staging of patients with hepatocellular carcinoma. Yoon et al. found a sensitivity of 100% for the detection of lung metastasis, lymph node metastasis and bone metastasis with PET, with specificities ranging from 84-100% [178]. However, the design of this study was unclear. Above this, imaging and clinical follow-up was used as reference standard. In a retrospective study including 121 patients with hepatocellular cancer, Ho et al. compared FDG-PET/CT with 11C-ACT-PET/CT and dual-tracer PET/CT [179]. Dual-tracer PET/CT was found to have a significantly better sensitivity than the comparators for the detection of distant metastasis (98% vs. 79% for FDG-PET/CT and 64% for ACT-PET/CT). This study suffered from incorporation bias.

Two prognostic studies were identified [180, 181]. One study with unclear design found a high SUV to be an unfavourable prognostic factor of overall survival [180]. Another prospective study found a high tumour to non-tumour SUV ratio to be an independent predictor of postoperative recurrence and overall survival [181].

Key messages

- **No new evidence was identified on the use of PET for the primary diagnosis of primary liver cancer (level 2).**
- **The evidence on staging (consisting of two small studies of low quality) is too limited to draw firm conclusions (level 2).**
- **PET can provide prognostic information in patients with primary liver cancer. However, this remains to be confirmed in good prospective studies.**

4.12 CERVICAL CANCER

4.12.1 Introduction

According to SIGN, the diagnosis of cervical cancer is made by the histopathological examination of cervical biopsies [182]. Cervical cancer is clinically staged using the so-called FIGO criteria. This FIGO staging does not take into account results of CT, MRI or PET. According to SIGN, all patients with visible, biopsy-proven cervical cancer should have MRI (or CT in case of contra-indications or clinically apparent FIGO stage IV) [182]. Staging with PET should be considered in patients not suitable for surgery [182].

In case of suspected recurrence, MRI or CT should be considered initially [182]. PET(/CT) should be considered in patients with proven recurrent disease on MRI or CT and with an option of salvage therapy [182].

Conclusion of the previous KCE report [1]:

- For staging, residual mass evaluation and recurrence diagnosis of cervical cancer, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.
- For recurrence diagnosis, there is inconclusive evidence that PET is superior to CT, because the specificity of PET is low compared with CT.

4.12.2 Diagnosis

The AHRQ 2008 report identified 1 primary study evaluating the diagnostic efficacy in a mixed population of patients with a primary diagnosis of cervical cancer (n = 75) and recurrent disease (n = 144) [56]. Since the data on detection of the primary tumour and local recurrence were mixed, and these were lesion-based, these will not be discussed here.

No additional primary studies were found.

4.12.3 Staging

Our search identified 1 HTA report of good quality [56], 1 systematic review of good quality [183] and 1 systematic review of moderate quality [114].

The AHRQ 2008 report identified 12 primary studies (published between 2003 and March 2008) on the use of PET for the initial staging of cervical cancer [56]. One of these studies combined the outcomes initial staging and staging of recurrent disease, another study combined the outcomes initial staging and detection of recurrence. In the 7 studies (n = 468) that used a patient-based analysis and that provided appropriate data, sensitivity ranged from 10-100% and specificity from 90-100% for lymph node staging [56]. One other small study reported a sensitivity and specificity of 100% for the staging of distant metastases.

The AHRQ 2008 report also identified 6 primary studies (published between 2003 and March 2008) on the use of PET/CT for the initial staging of cervical cancer [56]. One of these studies combined the outcomes initial staging and staging of recurrent disease. In the 2 studies that 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 (not calculable in the third study) for the detection of extra-cervical and/or metastatic disease [56].

Overall, when distinguishing pelvic lymph nodes from extrapelvic lymph nodes, a considerable amount of studies reported a low sensitivity for pelvic lymph node staging, but a moderate sensitivity for extrapelvic lymph node staging. Specificity was consistently high across both lymph node regions [56].

Selman et al. included 8 primary studies (published between 1999 and 2005) on the use of PET for the nodal staging of primary cervical cancer [183]. Two of these studies were also identified by the AHRQ 2008 report [56]. PET was found to have a significantly better pooled sensitivity (75%) and specificity (98%) than CT (32 studies; 58% and 92% respectively) and MRI (24 studies; 56% and 93% respectively), but sentinel node biopsy (31 studies) was found to be the most accurate diagnostic technique (pooled sensitivity and specificity of 91% and 100% respectively), in particular for early-stage disease. However, no direct comparisons were identified.

Bourguet et al. [114] identified 1 systematic review [184] and 2 additional primary studies (both included in the AHRQ 2008 report). In the included systematic review of Havrilesky et al., pooled sensitivity and specificity were 84% and 95% respectively for the detection para-aortic lymph node metastases, and 79% and 99% respectively for the detection of pelvic lymph node metastases [184]. Many of the primary studies included in Havrilesky et al. were also identified by Selman et al. [183].

No additional primary studies were identified by our search.

4.12.4 Detection of recurrence

The AHRQ 2008 report identified 13 primary studies (published between 2003 and March 2008) on the use of PET for the detection of recurrent disease [56]. One of these studies combined the outcomes initial staging and detection of recurrence. Most studies had a mixed reference standard (clinical follow-up or histology/biopsy). Only 5 studies provided appropriate data based on a per-patient-analysis. In the 2 studies that provided data on the diagnostic efficacy to detect recurrent disease considering all regions, sensitivity was 90% and 96% and specificity was 84% and 76% [56]. The 3 other studies provided data on 9 different regions separately (see table 14). Sensitivity ranged from 50% for bone and PLN sites to 100% for liver/spleen, MLN and ILN sites. Specificity ranged from 88% for MLN to 100% for liver/spleen, lung, PALN, PLN and ILN sites [56].

Table 14. Primary studies on diagnostic efficacy of PET for the detection of recurrent cervical cancer (per region) [56].

Region	Lin 2006		Yen 2006		Yen 2004	
	Se	Sp	Se	Sp	Se	Sp
Peritoneum	57%	89%	65%	98%	88%	96%
Bone	50%	96%	100%	97%	-	98%
Liver/spleen	100%	100%	67%	99%	100%	98%
Lung	75%	100%	92%	97%	78%	100%
Mediastinal lymph nodes	100%	88%	100%	96%	100%	98%
Supraclavicular lymph nodes	75%	95%	81%	98%	85%	98%
Para-aortic lymph nodes	90%	94%	88%	99%	88%	100%
Pelvic lymph nodes	50%	100%	83%	98%	91%	98%
Inguinal lymph nodes	100%	100%	83%	96%	100%	100%

The AHRQ 2008 report also identified 2 primary studies on the use of PET/CT for the detection of recurrent disease, involving a total of 64 patients [56]. The reference standard was histology/biopsy or clinical follow-up in both studies. Sensitivity was 83% and 90%, while specificity was 100% and 81%.

Bourguet et al. [114] identified 3 primary studies (already identified by the AHRQ 2008 report) and the systematic review of Havrilesky et al. [184]. Havrilesky et al. calculated a pooled sensitivity and specificity of 96% (95%CI 87-99%) and 81% (95%CI 58-94%) respectively for the detection of recurrence in patients with a suspicion of recurrence (3 primary studies). Pooled sensitivity and specificity for the detection of recurrence using systematic follow-up with PET were 92% (95%CI 77-98%) and 75% (95%CI 69-80%) respectively (2 primary studies).

Our search identified 1 prospective [185] and 2 retrospective primary studies [186, 187] (see table 15). The prospective study included 90 patients with a suspicion of recurrent uterine cancer, 50 of which had uterine cervical cancer [185]. The two retrospective studies only included patients with suspected recurrent cervical cancer [186, 187]. The 3 studies used histology/biopsy or clinical follow-up as reference standard, but 2 studies suffered from incorporation bias [185, 186]. For PET, sensitivity and specificity ranged from 80-92% and 74-93%, respectively [185-187]. For PET/CT, sensitivity was 91% and 92% in 2 studies, while specificity was 92% and 94% [185, 186].

Kitajima et al. found PET/CT to be significantly more sensitive than CT [185]. They also reported a change in treatment plan in 42% of patients with PET/CT.

Table 15. Primary studies on diagnostic value of PET, PET/CT or CT for the detection of recurrent cervical cancer.

Study ID	N	Design	PET		PET/CT		CT	
			Se	Sp	Se	Sp	Se	Sp
Kitajima 2009 [185]	90	P	80% (68-91%)	74% (61-87%)	91% (82-99%)	94% (86-100%)	68% (54-82%)	87% (77-97%)
Kitajima 2008 [186]	52	R	80% (64-95%)	78% (62-94%)	92% (81-100%)	93% (83-100%)	-	-
van der Veldt 2008 [187]	40	R	92% (81-96%)	93% (71-100%)	-	-	-	-

P = prospective, R = retrospective; Se = sensitivity, Sp = specificity. 95%CI are provided between brackets.

4.12.5 Staging of recurrence

The AHRQ 2008 report identified 3 primary studies on the use of PET(/CT) for the staging of recurrent cervical cancer [56]. The reference standard was histology/biopsy or clinical follow-up in all 3 studies. Two studies only reported a lesion-based analysis. The small prospective study that used a patient-based analysis reported a sensitivity and specificity of 92% and 100% respectively for the M-staging of recurrent cervical cancer [56].

Bourguet et al. identified 2 new primary studies, already identified by the AHRQ 2008 report [114].

Our search identified one additional primary study with retrospective design [187]. For the detection of local recurrence, a sensitivity and specificity of 100% and 97% respectively was found. For the detection of regional recurrence, a sensitivity and specificity of 87% and 100% was found. Sensitivity and specificity for M-staging were 75% and 100% respectively.

4.12.6 Prognosis

Bourguet et al. [114] identified one retrospective prognostic study reporting a better 5-year cause-specific survival in patients with normal FDG-uptake after external irradiation and intracavitary brachytherapy for cervical carcinoma compared to patients with abnormal FDG-uptake. A Cox proportional hazards model of survival outcome indicated that any abnormal post-therapy FDG-uptake was the most significant prognostic factor for developing metastatic disease and death from cervical cancer when compared with pre-treatment- and treatment-related prognostic factors ($p<0.0001$).

Our search identified 4 additional prospective prognostic studies, 2 using PET [188, 189] and 2 using PET/CT [190, 191] (table 16). All studies used multivariate analysis to detect independent significant prognostic factors. Different outcomes were used across the 4 studies. PET(/CT) lymph node status was found to be a prognostic factor in 2 studies [189, 191], while tumour volume and heterogeneity as determined by PET/CT was found to be a prognostic factor in another study [190]. In contrast to Kidd et al. [190], Xue et al. also found SUVmax to be a prognostic factor [189]. Other prognostic factors include SUVmax for para-aortic lymph nodes [188] and disease status on 3-months post-treatment PET/CT [191].

Table 16. Primary studies on prognostic value of PET and PET/CT.

Study ID	N	Index test	Design	Outcome	Prognostic PET(/CT)-parameter(s)
Kidd 2008 [190]	72	PET/CT	P	Recurrence OS	Tumour volume and heterogeneity, determined by PET/CT
Yen 2008 [188]	70	PET	P	RFS OS	SUVmax for para-aortic lymph nodes
Schwarz 2007 [191]	92	PET/CT	P	PFS OS	Progressive disease on 3-mo post-treatment PET/CT Partial metabolic response on 3-mo post-treatment PET/CT Pre-treatment PET/CT lymph node status
Xue 2006 [189]	96	PET	P	DFS OS	PET lymph node status SUVmax

P = prospective, R = retrospective; DFS = disease-free survival, OS = overall survival, RFS = recurrence-free survival, PFS = progression-free survival.

Key messages

- **A considerable amount 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).**
- **For the detection and staging of recurrence, new studies report sensitivities between 80 and 92% and specificities between 74 and 94%. The sensitivity of PET/CT seems to be superior to that of CT alone (level 2).**
- **PET and PET/CT have a prognostic value in patients with cervical cancer.**

4.13 OVARIAN CANCER

4.13.1 Introduction

The previous KCE report stated the following conclusions about the use of PET in the work-up of patients with ovarian cancer [1]:

- For diagnosis, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.
- For initial staging, there is no evidence.
- For diagnosis of recurrence, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.
- For evaluation of treatment response, there is no evidence.

4.13.2 Diagnosis

The AHRQ 2008 report identified 3 prospective studies evaluating the use of PET/CT for the primary diagnosis and staging of ovarian cancer (1 study) and the use of PET (1 study) and PET/CT (1 study) for the primary diagnosis of ovarian cancer [56]. All 3 studies used histology/biopsy as reference standard. For PET, sensitivity and specificity were found to be 78% and 86% respectively. For PET/CT, sensitivity was 87% and 100%, while specificity was 100 and 92% [56].

In 2006, the AHRQ already identified 3 primary studies on the use of PET for the evaluation of women with adnexal masses (of which one was included in the AHRQ 2008 report) [192]. Pooled sensitivity and specificity were found to be 67% (95%CI 52-79%) and 79% (95%CI 70-85) respectively. Of the evaluated imaging modalities (ultrasound, MRI, CT, PET), no evidence was found to support the superiority of any single modality, although PET appeared inferior to the rest.

Bourguet et al. identified one primary study that was already included in the AHRQ reports [114].

Our search identified one additional prospective study evaluating the use of PET/CT for the diagnosis of malignancy in women with suspected ovarian cancer [193]. Histopathology was used as reference standard. Sensitivity and specificity were 71% and 81% respectively.

4.13.3 Staging

The AHRQ 2008 report identified 4 primary studies, mostly with mixed outcomes [56]: one study on primary diagnosis and staging with PET/CT, two studies on primary staging alone with PET/CT (1 study) and PET (1 study), and one study on staging and detection of recurrence with PET. However, no results were found to be useful for our report.

Bourguet et al. also identified 4 primary studies (of which 2 were included in the AHRQ 2008 report) [114]. Again, none were found to be useful for the present report.

No additional primary studies were identified by our search

4.13.4 Detection of recurrence

The AHRQ 2008 report included 4 primary studies evaluating PET for the detection of recurrent ovarian cancer [56]. All studies used histology/biopsy or clinical follow-up as reference standard. Sensitivity ranged from 85-92%, while specificity ranged from 78-100%. In addition, 8 primary studies were included evaluating PET/CT for the detection of ovarian cancer [56]. Most studies used histology/biopsy or clinical follow-up as a mixed reference standard. Sensitivity ranged from 73-100%, while specificity ranged from 40-100%.

Bourguet et al. identified one systematic review and 3 additional primary studies (of which one was not included in the AHRQ 2008 report) [114]. The systematic review included 10 primary studies [184]. In the 5 studies evaluating the use of PET for the detection of recurrent disease in case of clinical suspicion, pooled sensitivity and specificity were 90% (95%CI 82-95%) and 86% (95%CI 67-96%) respectively.

In the 5 studies evaluating PET for the detection of recurrent disease during systematic follow-up, pooled sensitivity and specificity were 54% (95%CI 39-69%) and Sp 73% (56-87%) respectively.

Our search identified one additional prospective study comparing PET/CT and CT for the detection of recurrent disease in 132 patients with clinical suspicion [194]. A mixed reference standard of histopathology or clinical follow-up was used. The study suffered from incorporation bias. Contrast-enhanced PET/CT was found to be significantly more sensitive than CT (79% vs. 61%), specificity did not differ significantly (90% vs. 85%).

4.13.5 Staging of recurrent ovarian cancer

The AHRQ 2008 report identified 2 prospective studies evaluating the use of PET (1 study) and PET/CT (1 study) for the staging of recurrent ovarian cancer. However, both studies used a lesion-based analysis.

No additional primary studies were identified by our search.

4.13.6 Evaluation of treatment response

One prospective study identified by Bourguet et al. found a significant correlation between metabolic response after neoadjuvant chemotherapy and overall survival [114]. No additional prognostic studies were identified by our search.

Key messages

- **For the primary diagnosis of ovarian cancer, the results for PET are in line with the previous KCE report, which found evidence of diagnostic accuracy. For PET/CT, the results appear to be better. Therefore, PET(/CT) can be useful in doubtful cases (level 2).**
- **For the staging of ovarian cancer, no eligible studies were found to base recommendations on.**
- **For the detection of recurrent disease, new evidence shows somewhat better results than previously. In one new study, PET/CT was found to be more sensitive than CT alone. In case of clinical suspicion of recurrence, sensitivity is good, while specificity is moderate (level 2). In the absence of clinical suspicion, PET is not recommended due to low sensitivity and specificity (level 2).**
- **No evidence was found on the evaluation of treatment response.**

4.14 UTERINE CANCER

Bourguet et al. identified 2 primary studies on the use of PET in patients with endometrial cancer [114]. One prospective study evaluated PET for the detection and N-staging of recurrent disease. However, the authors only presented a lesion-based analysis. One small retrospective study evaluated PET for the detection of recurrent disease and the evaluation of treatment response. This study reported a sensitivity and specificity of 100% and 91% respectively [114].

Our search identified 4 additional primary studies. One small retrospective study compared PET with MRI for the assessment of myometrial infiltration in patients with clinical stage I uterine corpus cancer [195]. 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 evaluated the use of PET/CT for the N-staging of patients with primary endometrial cancer [196]. Histopathology was used as reference standard. Sensitivity and specificity were found to be 50% and 87% respectively.

Finally, two studies (one prospective and one retrospective) evaluated PET (1 study) and PET/CT (both studies) for the detection of recurrent uterine cancer [185, 197].

Both studies used histopathology or follow-up as reference standard, but suffered from incorporation bias. One of these studies included both patients with uterine cervical (n = 50) and endometrial cancer (n = 40) [185]. For PET/CT, sensitivity was 91% and 100%, while specificity was 94% and 95% [185, 197]. Kitajima et al. found PET/CT to be significantly more sensitive than CT [185].

Key messages

- **The evidence on the use of PET(/CT) is too limited to base recommendations on (level 2).**

4.15 RENAL CANCER

4.15.1 Introduction

The early diagnosis of renal carcinoma may remain clinically occult for a large part of its course. Traditional symptoms include pain, haematuria, and occurrence of flank mass, but these perceptible symptoms are present in only 9% of renal cancer patients and are often indicative of advanced disease. A tumour in the kidney can insidiously progress, swelling in the retroperitoneum before metastases appear. Approximately 30% of patients with renal carcinoma present with metastases, 25% with locally advanced renal carcinoma and 45% with localized disease [198].

The ability to accurately detect and characterize renal masses and to stage malignant renal tumours is crucial for the management of patients. An accurate diagnosis and staging will help oncologists and surgeons to choose the more appropriate therapeutic approach (e.g., surgery or systemic treatment). For example, the early detection of a tumour can make minimally invasive surgery and partial nephrectomy possible instead of standard radical nephrectomy.

Conventional imaging includes CT, ultrasonography and MRI [56]. However, until now, morphological imaging methods present several diagnostic problems in differentiating between benign and malignant solid renal tumours, as well as in evaluating tumour spread and distant disease [56]. The previous KCE report found only few studies on PET [1]. For initial diagnosis and detection of recurrence, there was a lack of evidence for diagnostic accuracy. For staging, evidence of diagnostic accuracy was found.

One new HTA report [56] and one systematic review [199] were identified. Additionally, one retrospective study compared the diagnostic performance of PET/CT and conventional imaging (plain chest radiography, abdominopelvic CT and whole body bone scan) [200].

Conclusion of the previous KCE report [1]:

- For initial diagnosis and detection of recurrence, there is a lack of evidence for diagnostic accuracy.
- For staging, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.

4.15.2 Diagnosis

The systematic review conducted by the AHRQ [56] included 1 study for primary diagnosis only and 4 studies for combined diagnosis and staging. The systematic review conducted by Bourguet et al. [199] only included two retrospective studies focusing on diagnosing and staging (both included in the AHRQ 2008 report).

The five primary studies included 2 prospective and 3 retrospective studies [56]. A total of 177 patients was included with sample sizes ranging from 15 to 66 patients. In 3 studies, the reference standard was histology/biopsy or clinical follow-up (3-6 months), whereas in the 2 remaining studies histology/biopsy was used exclusively as reference standard. In the two prospective studies, sensitivity was 87% and 47% and specificity was 75% and 80% for the detection of the primary tumour. AHRQ conducted a meta-analysis on the 3 retrospective studies and another meta-analysis on the two retrospective studies that used histology/biopsy as reference standard. However, only pooled negative and pooled positive likelihood ratios were provided.

Pooled sensitivity and specificity were not reported. Across these five primary studies sensitivity ranged from 47-90%, while specificity ranged from 75-100% [56]. Bourguet et al. [199] synthesized the diagnostic performance of PET scan compared to other imaging techniques in detecting primary tumours or metastases (see table 17). Authors concluded that although PET imaging was more specific than conventional imaging, its use was limited by its low sensitivity.

Table 17. Comparison of PET with other imaging techniques in detecting primary tumours or metastases.

Accuracy of FDG-PET vs abdominal CT: primary RCC tumours	
FDG-PET	**Abdominal CT**
Se: 60.0%	Se: 91.7%
Sp: 100%	Sp: 100%
Accuracy of FDG-PET vs abdominal CT: retroperitoneal lymph node metastases and/or renal bed recurrence	
FDG-PET	**Abdominal CT**
Se: 75.0%	Se: 92.6%
Sp: 100.0%	Sp: 98.1%
Accuracy of FDG-PET vs chest CT : metastases to the lung parenchyma	
FDG-PET	**Chest CT**
Se: 75.0%	Se: 91.1%
Sp: 97.1%	Sp: 73.1%
Accuracy of FDG-PET vs chest CT+ bone scan: bone metastases	
FDG-PET	**Chest CT+ bone scan**
Se: 77.3%	Se: 93.8%
Sp: 100.0%	Sp: 87.2%

4.15.3 Staging

The AHRQ 2008 report [56] identified one small prospective study on staging and two retrospective studies that included patients for diagnosis, staging and restaging. The prospective study included 24 patients with histologically proven renal cell carcinoma with metastatic disease awaiting a therapeutic decision for surgery, radiofrequency ablation, general specific treatment (immunotherapy) before surgery, or monitoring. PET findings resulted in five changes (21%) to the management strategy: from observation to surgery (n = 2) or immunotherapy (n = 2) and from surgery to immunotherapy (n = 1).

4.15.4 Restaging

The AHRQ 2008 report identified one retrospective study evaluating the use of PET for the restaging of patients with RCC who had undergone nephrectomy [56]. PET detected 64% of all soft tissue metastasis and 79% of bone metastasis. According to the localisation of metastases (lymph nodes, lung, liver, bone), PET demonstrated a lower sensitivity (50-75%) than conventional imaging (77-100%), but a higher specificity (97-100% vs. 73-98%).

4.15.5 Detection of recurrence

Bourguet et al. [199] identified 4 primary studies, all having a retrospective design, conducted to detect recurrences after treatment. The AHRQ also reported results for the same studies in its appendices without specific interpretation of results for recurrence [56]. In the 2 trials specifically reporting on detection of recurrence, sensitivity was 63% and 71%, while specificity was 100% and 75%. Due to the small sample sizes, confidence intervals were very large and no firm conclusion can be drawn.

Our literature search identified 1 additional retrospective study [200] which compared diagnostic performance of PET/CT and conventional imaging (plain chest radiography, abdominopelvic CT, whole body bone scan) to detect recurrences in 63 patients with RCC followed after surgical treatment. PET/CT had similar sensitivity (94% vs. 97%) and specificity (84% vs. 81%) as conventional imaging.

4.15.6 Monitoring of treatment response

No HTA reports or systematic reviews were identified that addressed or found studies on this indication. Our additional search identified no primary study that focused on this purpose.

Key messages

- **The evidence on initial diagnosis and staging is limited to small studies of low quality reporting wide confidence intervals. For the initial staging, PET can be useful when CT and/or bone scan are equivocal, although this remains to be confirmed by good prospective trials (level 2).**
- **One small retrospective study on restaging provides insufficient information to base recommendations on (level 2).**
- **For the detection of recurrent disease, PET(/CT) has a moderate diagnostic efficacy that is similar to that of CT alone. Therefore, it can be considered if CT is equivocal (level 2).**

4.16 TESTICULAR CANCER

4.16.1 Introduction

The previous KCE report stated the following conclusions about the use of PET in the work-up of patients with testicular cancer [1]:

- For staging and residual mass detection, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.
- For therapeutic response and detection of occult recurrence, there is a lack of evidence for the use of PET.

4.16.2 Staging

Both the AHRQ 2008 report [56] and Bourguet et al. [201] identified one prospective study evaluating the use of PET for the staging of 46 patients having undergone orchidectomy and negative postoperative conventional staging. 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.

Our search identified 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 [202]. 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.

4.16.3 Detection of recurrence

The AHRQ 2008 report identified 2 small primary studies (1 prospective, 1 retrospective) evaluating the use of PET for the detection of recurrent testicular cancer [56]. The prospective study used histology/biopsy as reference standard, while the retrospective study used clinical follow-up. Sensitivity was 100% in both studies, while specificity was 47% and 72%.

No additional primary studies were identified by our search.

4.16.4 Evaluation of residual mass

The AHRQ 2008 report identified one prospective study evaluating the use of PET for the evaluation of residual testicular masses [56]. However, this study used a lesion-based analysis.

Bourguet et al. identified 2 primary studies (one also identified by the AHRQ) [201]. The study that used a patient-based analysis found a sensitivity and specificity of 62% and 83% respectively, which was not significantly better than CT/MRI.

Our search identified one additional prospective study evaluating PET for the evaluation of residual masses in 121 patients with non-seminomatous germ cell tumours and a primary or metastatic retroperitoneal tumour of at least 5 cm or with distant metastases at the time of primary diagnosis or first relapse [203]. Histology was used as reference standard. Sensitivity and specificity were found to be 70% and 39% respectively.

Key messages

- **The new evidence on the use of PET(/CT) for the staging of testicular cancer consists of 2 primary studies and is in line with the previous report (level 2). Overall, this evidence remains inconclusive.**
- **New studies confirm that the diagnostic efficacy for residual mass evaluation is only moderate. Sensitivity is not higher than 70%, one study showed a low specificity of 39% (level 2).**
- **For the detection of recurrent disease, two small studies showed a good sensitivity but with wide confidence intervals. Therefore, firm conclusions cannot be drawn (level 2).**
- **For the prediction of treatment response, no new evidence was identified.**

4.17 PROSTATE CANCER

The AHRQ 2008 report identified 4 primary studies evaluating PET (and PET/CT in 1 study) for the detection and/or staging of recurrent disease [56]. Different reference standards were used across these studies. One of these studies suffered from partial verification. Sensitivity ranged from 32-75%, while specificity ranged from 0-100% [56].

Bourguet et al. [201] identified one small primary study in addition to those identified by the AHRQ 2008 report. This study evaluated the use of PET for the monitoring of treatment response in 23 patients with metastatic prostate cancer treated with chemotherapy. PET correctly identified 85% of the responders.

No additional primary studies were identified by our search.

Key messages

- **For detection of recurrent prostate cancer, sensitivity appears to be too low. Therefore, it is not recommended to use PET(/CT) for this indication (level 2).**
- **The evidence on the use of PET(/CT) for other indications is too limited to draw firm conclusions (level 2).**

4.18 BLADDER CANCER

The AHRQ 2008 report identified 2 prospective studies on the use of PET for the primary staging of bladder cancer [56]. Sensitivity was 53% and 77%, specificity was 72% and 94%. One other retrospective study evaluated the use of PET and PET/CT for the staging of recurrent disease [56]. Sensitivity and specificity were 90% and 85% respectively.

No additional primary studies were identified by our search.

Key messages

- **The evidence on the use of PET(/CT) is too limited to base recommendations on (level 2).**

4.19 PENILE CANCER

No eligible studies were identified on the use of PET(/CT) in patients with penile cancer.

Key messages

- **No eligible studies were identified on the use of PET(/CT) in patients with penile cancer.**

4.20 GASTROINTESTINAL STROMAL TUMOURS

No evidence from HTA reports or systematic reviews was found to update the conclusions formulated in the previous KCE report [1]. For diagnosis, there is no evidence for the use of PET. For therapy monitoring, there is a potential impact of PET on therapy planning.

Key messages

- **No new evidence was identified since the previous KCE report. The recommendations therefore remain unchanged (level 2). PET has a potential impact on therapy planning.**

4.21 BRAIN CANCER

4.21.1 Introduction

Specific symptoms of brain tumours depend on the tumour's size, location, degree of invasion, and related swelling. Headaches, seizures, weakness in one part of the body, and changes in the person's mental functions are most common. Diagnosis starts with a complete medical history, physical examination and a careful neurological assessment, followed with fundoscopy and a focused neurologic examination. Appropriate brain imaging and histopathology are required to confirm diagnosis [56].

The initial screening of brain tumours is done by MRI, which produces higher resolution images and can access more areas of the brain than CT. MRI is also used for neurosurgical planning and risk assessment [56]. To distinguish infiltrative brain tumours from non-neoplastic conditions, high-grade from low-grade tumours, and primary tumours from metastastic tumours, magnetic resonance spectroscopy (MRS) may be used [56].

The reference diagnostic standard remains tissue biopsy. More recently developed stereotactic biopsy techniques are minimally invasive, with decreased morbidity and mortality relative to traditional neurosurgery. Stereotactic biopsy should be obtained to help confirm diagnosis of low-grade gliomas [56].

MRI, MRS and PET can assist in tumour localization for biopsy. Testing for biomarkers may also assist in diagnosis, treatment planning and predicting prognosis [56].

Conclusion of the previous KCE report [1]:

- For diagnosis, i.e. distinguishing high-grade from low-grade glioma, there is evidence of diagnostic accuracy including the determination of sensitivity and specificity.
- For diagnosis, i.e. biopsy targeting and delineation of lesion for therapy planning, there is some evidence of diagnostic accuracy.
- For restaging, i.e. distinguishing recurrent malignancy from radiation necrosis, there is some evidence of diagnostic accuracy.

4.21.2 Diagnosis

The AHRQ 2008 report [56] assessed the diagnostic efficacy of PET in the management of brain tumours. The systematic review identified one small prospective study that used PET to diagnose brain tumours in 30 patients (both primary diagnosis and recurrences, for low- and high-grade tumours). In this study, the reference standard was either histology/biopsy or clinical follow-up (mean 20 months). For the combined outcomes (primary diagnosis and recurrences), PET was found to be insufficiently accurate to be recommended for diagnosis of brain cancer. Using FDG, sensitivity was only 61% (95%CI 39-79%) and specificity was 43% (95%CI 12-80%). Using FDOPA, sensitivity was higher 96% (95%CI 76-99%), but specificity still was low 43% (95%CI 12-80%).

Our own search identified two additional studies [204, 205]. Pöpperl et al. conducted a diagnostic accuracy study among 54 adult patients with suspected supratentorial primary gliomas [204]. FET-PET was evaluated as index test and histopathology or clinical follow-up served as reference standards.

FET-PET was able to differentiate low-grade from high-grade tumours with a high sensitivity 93% (95%CI 76-99%) and a high specificity 100% (95%CI 80-100%). The authors concluded that histopathologic examinations remained the gold standard for establishing tumour grade. However, dynamic FET-uptake evaluations contributed significantly to predicting ultimate histological findings.

Roessler et al. conducted a prospective study with MET-PET for the detection and grading of brain tumours among 27 patients with suspected cerebral gliomas [205]. In this group, all patients were found to have a brain tumour. MET-PET had a sensitivity of 96% (95%CI 79-99%) for the diagnosis of brain tumours. For the differentiation between low-grade and high-grade tumours, no sensitivity and specificity could be calculated.

4.21.3 Staging

The AHRQ 2008 report [56] identified three small studies using PET to stage patients with suspected primary glioma (2 studies) or patients with primary astrocytomas (1 study). Two studies were prospective. All studies used histology/biopsy as reference standard. In the 2 studies where a 2x2 table was provided, sensitivity was only 63% and 75% and specificity was 100% and 0% [56].

4.21.4 Detection of recurrence

The AHRQ 2008 report [56] identified only one small study that used PET to detect recurrences in 28 patients with glioblastoma multiforme after surgical and/or conservative treatment (mean follow-up 13 months). This retrospective study used MRI and MET-PET (carbon-11 methionine PET) as reference standards. For the detection of recurrences, PET had a low sensitivity (11%; 95%CI 2-36%). Even when PET was compared to survival longer than 12 months, sensitivity (7%; 95%CI 0.4-38%) as well as specificity (14%; 95%CI 2.5-44%) remained unacceptable. Moreover, PET did not seem to be highly discriminative in identifying the stage of the disease, and in distinguishing between necrosis and recurrences. These limits hampered the authors to draw firm conclusions about the diagnostic utility of PET for brain cancer.

Our own search identified three additional studies [206-208]. One retrospective study [206] included 38 patients referred with possible brain tumour recurrence (n=32) or newly diagnosed brain tumour (n=6). Reference standards included histopathology (n=21) and clinical follow-up (n=17). FDG-PET was found to have a sensitivity of 74% (95%CI 53-83%) and specificity of 73% (95%CI 39-86%). In the subset of patients where only histopathology was considered (n=21), sensitivity was quite similar (72%; 95%CI 47-89%), but specificity was lower (33%; 95%CI 2-87%). Pöpperl et al. [207] conducted a study with FET-PET for the detection of brain tumour recurrence after locoregional radio-immunotherapy among 24 patients with proven malignant gliomas (5 anaplastic astrocytomas, 19 glioblastomas). Among the 17 of 24 patients who presented with tumour progression, 10 had tumour recurrence and 7 had re-growth of residual tumour. FET-PET yielded a high sensitivity of 94% (95%CI 69-99%), but a moderate specificity of 71% (95%CI 30%-95%) to detect recurrence.

FET-PET demonstrated the highest discrimination capacity between patients with tumour recurrence and tumour-free patients at a threshold value of 2.4 for the TUmax/BG ratio. Finally, Rachinger et al. [208] included 45 consecutive patients with gliomas to test the ability of FET-PET to detect tumour recurrence or tumour progression after treatment. In patients who already had suspected recurrent tumour on the basis of MRI, FET-PET yielded a high sensitivity of 100% (95%CI 86-100%) and a specificity of 93% (95%CI 64-100%). The authors concluded that FET-PET is useful to differentiate side effects of therapy from tumour recurrence.

4.21.5 Restaging

No HTA-reports or systematic reviews were found on this indication.

4.21.6 Monitoring of treatment response

No HTA-reports or systematic reviews were found on this indication.

4.21.7 Prognosis

Two primary studies were retrieved by our own literature search [209, 210].

Spence et al. [210] followed 22 patients with glioblastoma multiforme (GBM) to determine overall survival and time-to-progression (TTP). The prognostic variables considered were FMISO hypoxic volume (HV) and tissue to blood concentration (T/B) max; age; Karnofsky performance status (KPS); extent of resection; and the MRI volumes, TO, TI Gd, T2, and T2-TO. Multivariate analyses for survival and TTP against the covariates HV (or T/B max), MRI T1Gd volume, age, and KPS reached significance only for HV (or T/B max; $p<0.03$). The multivariate model indicated that an increase in HV in the order of 7 to 8 cm3 was associated with a 50% reduction in survival or TTP. An increase in T/B max of 0.25 (or 0.35 for TTP) was associated with a 50% reduction in survival (or TTP).

Ceyssens et al. [209] conducted a prognostic study among 52 patients with MET-PET. Overall median survival was 34.9 months. In a proportional hazard Cox regression model with age, WHO grading, and MET-uptake index, only WHO grading was significantly predictive of survival ($p = 0.015$), whereas age and MET-uptake index were not significantly predictive. Moreover, no thresholds could be found at which MET could be considered predictive of survival (Kaplan-Meier statistics).

Key messages

- **For the diagnosis of brain cancer, newly identified evidence does not change the previous recommendations (level 2). There is evidence of diagnostic accuracy for distinguishing high-grade from low-grade glioma and for biopsy targeting and delineation of the lesion for therapy planning.**
- **FDG-PET scanning is insufficiently accurate to be recommended for staging of brain cancer (level 2).**
- **For the detection of recurrent disease, new evidence is in line with the previous KCE report, showing low diagnostic accuracy for FDG-PET (level 2). Results with amino acid-PET appear to be better (level 2).**
- **PET can provide prognostic information in patients with brain tumours, although the clinical consequences are unclear.**

5 PET IN CARDIOLOGY

5.1 MYOCARDIAL PERFUSION

The previous KCE report did not identify evidence for the use of PET for myocardial perfusion evaluation [1].

Our search identified 2 systematic reviews evaluating PET (mainly with the tracers Rubidium-82 and Nitrogen-13) for the evaluation of myocardial perfusion in patients with suspected coronary artery disease. Nandalur et al. identified 19 primary studies involving 1442 patients [211]. Reference standard was catheter x-ray angiography (≥50% diameter stenosis as threshold for significant CAD). Patient-based pooled sensitivity and specificity were 92% and 85%, respectively. However, the included studies showed an overall low quality. Beanlands et al. identified 14 primary studies (1460 patients) [212], of which 13 were also included in the systematic review of Nandalur et al. Sensitivity and specificity of PET ranged from 83-100% and 73-100%, respectively. These figures were similar for multislice CT (n=19; Se 85-100%, Sp 67-98%) and dobutamine stress MR (n=8; Se 86-96%, Sp 80-86%), but slightly worse for MR angiography (n=31; Se 38-90%, Sp 73-100%). According to the external experts, it should be stressed that the comparison between PET (i.e. a functional imaging technique, and if combined with CT also a morphological technique) and purely morphological techniques (multislice CT, MR angiography) is difficult, and that these techniques should be considered complementary.

No additional diagnostic accuracy studies were identified.

5.2 MYOCARDIAL VIABILITY

The previous KCE report found evidence of diagnostic efficacy up to diagnostic thinking to select patients eligible for revascularisation [1].

Our search identified the MAS 2005 report [213], being an update of the ICES HTA 2001, which was included in the previous KCE report [1]. Nine new primary studies were identified (published before April 2005). Sensitivity of PET for the evaluation of myocardial viability ranged from 75-100%, specificity ranged from 76-100% [213]. PET was found to have a comparable diagnostic accuracy as dobutamine echocardiography, although dobutamine echocardiography had a better specificity. Beanlands et al. also found PET to have a comparable sensitivity (range 80-100%) as dobutamine stress MR (range 77-100%) and late gadolinium enhancement MR (range 72-98%), but a lower specificity [212].

Our search identified one RCT randomising 430 patients with severe left ventricular dysfunction and suspected coronary disease to management assisted by FDG-PET or standard care [214]. At 1 year, no significant differences were found between the study groups as to the composite outcome of cardiac death, myocardial infarction or recurrent hospital stay for cardiac cause (HR PET vs. standard care: 0.78; 95%CI 0.58-1.1, p=0.15). However, this study received criticism on the fact that only 75% of the patients who underwent PET imaging adhered to the PET recommendation. In a post-hoc analysis, which compared the 'adherence to PET' subgroup with standard care, a HR of 0.62 (95%CI 0.42-0.93, p=0.019) was found [214].

5.3 PROGNOSIS

Beanlands et al. included 5 prognostic studies, reporting hard cardiac event rates in 0.09-0.9% of patients with proven coronary artery disease and normal PET results vs. 7% of patients with abnormal PET results [212]. However, it is not always clear whether these studies used multivariate analysis. The better studies had sample sizes between 367 and 629 patients.

Our search identified 2 additional prognostic studies. Tio et al. found the myocardial perfusion rate measured by FDG-PET to be an independent predictor of cardiac death (HR 4.11; 95%CI 2.98-5.67) in 480 patients with advanced ischemic disease [215]. On the other hand, Santana et al. did not find 82Rb/gated FDG-PET to be predictive of cardiac death in 104 patients with ischemic cardiomyopathy [216].

One of the consulted experts provided one additional prospective prognostic study including 261 patients with ischemic cardiomyopathy undergoing PET (with NH3 and FDG) for the assessment of myocardial viability (ref). In the 167 patients not undergoing revascularisation, multivariate analysis found age (HR 2.1, 95%CI 1.2-3.7), presence of left bundle branch block (HR 3.4, 95%CI 1.6-7.2) and extent of perfusion-metabolism mismatch on PET (HR 1.36, 95%CI 1.1-1.6) to be predictive of cardiac death during a median follow-up period of 2.1 years.

Key messages

- **Meta-analyses show moderate to good diagnostic efficacy of PET for the evaluation of myocardial perfusion (level 2).**
- **For the evaluation of myocardial viability, diagnostic accuracy studies show moderate to good diagnostic efficacy of PET that is comparable to other techniques (level 2).**
- **New studies suggest a prognostic value of PET in patients with ischemic heart disease.**

6 PET IN NEUROLOGY

6.1 PARKINSON DISEASE

6.1.1 Introduction

Parkinson disease (PD) is a degenerative brain disorder characterized by the presence of Lewy bodies and a degeneration of dopaminergic neurons in the substantia nigra, with loss of their nerve terminals in the basal ganglia structures (striatum). It has an overall estimated prevalence of up to 0.3%; but it is more common at older age affecting as many as 0.5-1% of individuals aged 65-69 years and 1-3% of individuals older than 80 years [217, 218]. Although PD is common, it can be difficult to diagnose clinically, particularly in early stages. In 1992, Hughes et al. [219] showed that up to 20% of patients diagnosed with PD reveal alternative diagnoses at autopsy. Based on these studies, clinical diagnostic criteria have been developed, which are now widely used in daily practice. New neuropathological data in 2001 suggested that an accuracy of 90% is the best that can be achieved with clinical assessment and clinical diagnostic criteria [217]. However, these numbers are largely drawn from clinical data of PD patients coming to the end stage of the disease, when clinical features and disease course are much more informative than in early stages. Accuracy in early stage patients is probably less [5].

Whereas classical MRI is a simple and relatively inexpensive technique that is widely available, a significant atrophy of the basal ganglia only becomes obvious in advanced disease stages. Nevertheless, MRI can be useful in the differential diagnosis of PD and symptomatic parkinsonism, e.g. vascular insults (or "vascular" parkinsonism), normal pressure hydrocephalus or degenerative disorders with pathognomonic MRI changes [220].

However, the clinical differentiation between typical idiopathic Parkinson's disease and atypical parkinsonian disorders remains complicated by the presence of signs and symptoms common to both forms of parkinsonism. Atypical parkinsonian disorders (e.g. multiple system atrophy, progressive supranuclear palsy, corticobasal degeneration, akinetic rigid features in dementia with Lewy bodies) usually have a different natural clinical course and tend to respond less well to dopaminergic treatment, making a precise diagnosis important for management as well as for prognostic reasons [221]. The sensitivity of the published clinical diagnostic inclusion criteria as compared to post-mortem neuropathology for these disorders amounts to 86% for multiple system atrophy [222] and 78% for progressive supranuclear palsy [223].

Functional imaging by means of SPECT and PET might be capable to distinguish between PD and other forms of parkinsonism, as these techniques depict the loss of neurotransmitter function typical for this disease. As compared to PET, SPECT is assumed to be less sensitive but more widely available because of the longer half-life of its radiotracers [220]. However, a full appraisal of the diagnostic value of SPECT to discriminate between PD and other disorders by the use of presynaptic (e.g. 123-I-bèta-CIT, 123-I-FP-CIT, 99-Tc-TRODAT-1) as well as postsynaptic tracers (e.g. 123I-BZM) is beyond the scope of this report; the interested reader is referred to a recent meta-analysis on this subject [224].

Currently, metabolic brain imaging with FDG and PET has been described as potentially useful in differentiating idiopathic PD from atypical forms. Brain receptor binding ligands are another type of radiotracers that are proposed for the same purpose. They include two distinct categories: the presynaptic ligands, such as 18F-Dopa, 11C-dihydrotetrabenazine, 11C-CFT, 18F-CFT etc., and the post-synaptic or D2 receptor radioligands, such as 11C-raclopride. While the presynaptic ligands theoretically have the potential to discriminate between PD and other neurological disorders, such as essential tremor or Alzheimer's disease, the postsynaptic ligands are assumed to allow for discrimination between PD and atypical parkinsonian disorders. Especially the postsynaptic ligand 11C-raclopride has a short half-life of 20 minutes, limiting it's applicability for routine diagnostic purposes.Besides diagnostic evaluation in individual patients with unclear PD diagnosis or prognosis of PD patients, PET scan has also been used in studies evaluating neuroprotection by drugs and detection of pre-clinical PD.

However, none of these applications has come yet to a stage where its use in routine clinical management of individual patients has been studied. Therefore, these applications are considered to be beyond the scope of this report.

6.1.2 Systematic reviews and meta-analyses

One good-quality systematic review on diagnosis and prognosis in Parkinson disease was identified [217]. This publication, reviewing the literature up to 2004, describes the results of one diagnostic case-control study, and could not retain any prognostic study of sufficient quality. It is concluded that there is insufficient evidence to support or refute PET scan (FDG-PET as well as PET and presynaptic/postsynaptic radioligands, e.g. FDOPA-PET) for diagnosing PD or for the prediction of disease progression in PD.

6.1.3 Primary studies

An additional search identified 3 potentially eligible diagnostic accuracy studies [7, 221, 225]. However, quality appraisal revealed an invalid reference standard in 2 studies (i.e. clinical diagnosis, where this clinical diagnosis was already unclear at inclusion) [7, 225] and partial verification in the third study (patients without clear clinical diagnosis after 2.1 years of follow-up were excluded [221]. Therefore, no primary studies were included ultimately.

No prognostic studies of sufficient quality are currently available.

Key message

- **A systematic review and an additional search for primary studies could not identify eligible studies on the use of PET for the diagnosis of Parkinson disease.**

6.2 ALZHEIMER'S DISEASE

6.2.1 Introduction

In 2005, the proportion of people with dementia in Belgium was estimated at 1.22% to 1.35%, corresponding to 127 174 and 140 639 subjects. For subjects 65 years of age and older, the published range of prevalence ranges from 6.3% to 9.3%. Alzheimer's disease (AD) or Alzheimer's dementia accounts for more than half of the cases of dementia (around 60%) [226]. Memory impairment is usually one of the first characteristics of AD. As the disease progresses cognitive deficits start to interfere with usual activities. The reference standards for the clinical diagnosis of AD are the criteria defined in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV-TR) and the NINCDS-ADRDA criteria (National Institute of Neurological Disorders and Stroke – Alzheimer Disease and Related Disorders). With the exception of some genetically well-defined forms of AD, the gold standard for the diagnosis of AD still remains the post-mortem confirmation by histopathology [226]. A potential usefulness of PET in the early diagnosis of AD may be related to its ability to demonstrate a reduced glucose metabolism (FDG-PET) in certain areas of the brain.

The previous KCE report [1] reviewed 4 HTA reports and 2 systematic reviews on the role of PET in the diagnosis of AD. It was concluded from these studies that there is evidence of diagnostic accuracy for AD. The sensitivity of FDG-PET scan in diagnosing AD compared to normal subjects was estimated at 88% (95%CI 79-94%) and its specificity at 87% (95%CI 77-93%). The sensitivity of PET in diagnosing AD among other forms of dementia varied from 86% to 95% and its specificity from 61% to 73%. However, it was recommended to treat these results cautiously because of methodological limitations of the included primary studies.

The conclusion of the included HTA reports and systematic reviews was that the routine use of FDG-PET scan in the diagnosis of AD could not be recommended, and that treatment without further testing was superior to treatment based on additional test using PET due to the limited effectiveness of pharmacological treatment of AD.

6.2.2 Systematic reviews and meta-analyses

A recent KCE report (in publication), specifically dealing with Alzheimer's disease, provides a recent literature update (up to June 2008) on this subject and includes systematic reviews, meta-analyses and HTA reports based on primary studies up to 2006.

Three large reviews were retained: a systematic review by SBU [227], by the EFNS task force [228] and by NICE-SCIE [226].

According to SBU, FDG-PET has moderate value in differentiating AD from normal subjects and from other dementia disorders [227]. This conclusion was based on six primary studies, all conducted in specialised services. Three studies used histopathological confirmation as reference standard. Sensitivity and specificity were both more than 80% in 3 of the 6 studies. SBU also concluded that it is not obvious that PET is superior to SPECT in differentiating AD from normal subjects or from other dementia disorders, since the likelihood ratios (LR+ varying from 2.5 to 13 for PET and from 2.5 to 16 for SPECT) were similar regardless of whether PET or SPECT was used. None of the included studies directly compared PET and SPECT.

According to EFNS, PET may be useful in those cases where diagnostic uncertainty remains after clinical and structural imaging work up, and should not be used as the only imaging measure [228]. Its literature review mainly relies on one of the systematic reviews already included in the previous KCE report.

NICE-SCIE confirms the diagnostic value of FDG-PET in AD, also mainly relying on the same systematic review [226]. The conclusion from this agency is that PET can yield a sensitivity of around 90% and a specificity of around 70%, may show some superiority over perfusion SPECT scan in detecting AD, but remains an expensive and invasive investigation.

An additional search for the present report yielded one moderate-quality meta-analysis [229]. Yuan et al. included 24 primary studies on the use of FDG-PET, SPECT and structural MRI in patients with mild cognitive impairment. None of these studies evaluated more than one diagnostic technique, making direct comparison of the different techniques difficult. The results of this meta-analysis are further limited by heterogeneity between the included studies and by a marked asymmetry in the funnel-plot, suggesting publication bias. Taking these limitations into consideration, FDG-PET seems to have a moderate diagnostic efficacy to detect AD in patients with mild cognitive impairment, with a pooled sensitivity of 89% (95%CI 82-94%) and a specificity of 57% (95%CI 78-90%).

6.2.3 Primary studies

No primary studies were identified directly comparing PET to other diagnostic tools (e.g. SPECT or clinical evaluation) and corresponding to the preset inclusion criteria.

Recently, PET scan tests demonstrating AD brain amyloid deposits (one of the three histopathological hallmarks of AD) have been developed, but these promising new tools deserve further diagnostic evaluation.

Recommendations

- **New evidence from one systematic review confirms the conclusions of the previous KCE report, with good sensitivity and low to moderate specificity for the diagnosis of Alzheimer's disease in patients with dementia (level 2).**
- **Although the results were heterogeneous, PET seems to have a moderate diagnostic efficacy to detect Alzheimer's disease in patients with mild cognitive impairment (level 2). However, there is still a need for a precise specification of the subgroup of patients for whom there will be an impact on clinical decision making.**

6.3 EPILEPSY

6.3.1 Introduction

Epilepsy is a common chronic disorder that affects approximately 3% of the population during their life-time. About 60-70% of patients experience focal (localization-related) or partial seizures, and 30-40% generalized seizures. Epilepsy is controlled with medication in approximately 70% of cases. Of the 20-30% of people who continue to have seizures despite drug treatment, the majority has a focal epilepsy of which temporal lobe epilepsy represents the most common form in adults [230, 231]. The precondition that a potentially surgically remediable epileptic syndrome should present a unique epileptogenic zone not overlapping with eloquent brain regions, has recently been expanded to some other forms of epilepsy. The proportion of drug-resistant patients who could or should be offered a surgical treatment remains unknown. Estimates vary from 5 to 30% of all drug-resistant epilepsies, with an annual need for surgery of around 1.5% [230].

Conclusion of the previous KCE report [1]:

- For pre-surgical evaluation of refractory epilepsy, there is some evidence of diagnostic accuracy but the added clinical value of PET is unclear. However, this is a rare indication.

6.3.2 Systematic reviews and meta-analyses

One meta-analysis on the use of PET in patients undergoing epilepsy surgery was identified [232], including 42 primary studies published before 2004 and 4 studies published in 2004-2006. However, because no systematic quality appraisal was performed on the included publications, the meta-analysis was excluded from our review. Another systematic review published by the Ontario Medical Advisory Secretariat in 2006 [233] also had to be excluded for the same reason.

One other good-quality systematic review was identified [230], reporting a literature search up to December 2003. This review largely covered the same time period as the HTA-MSAC 2004 already included in the KCE 2005 report [1]. However, it additionally compared PET and other imaging techniques, such as CT scan, conventional MRI, volumetric MRI, ictal and interictal single photon emission computed tomography (SPECT), and magnetic resonance spectroscopy (MRS).

Whiting et al. included 19 studies on the diagnostic accuracy of PET in adults or children with refractory epilepsy being considered for surgery [230]. Most studies described results for FDG-PET, four studies additionally included other radioligands (tracers). A significant heterogeneity was found between the studies, possibly due to differences in study design, study population, index test characteristics or reference standard(s). Because of this heterogeneity, the results could not be pooled. The following reference standards were accepted in the review: ictal EEG, a combination of tests (usually ictal EEG, neuropsychological testing, Wada testing and other imaging techniques), or site of surgery. A separate analysis was performed to evaluate the prognostic value of PET, i.e. its capacity to predict surgical outcome.

Taking the study heterogeneity into consideration, PET seems to perform equally well as ictal SPECT (7 studies) in localising the epileptogenic focus [230]. It seems to perform better than interictal SPECT (4 studies). When compared to conventional MRI (7 studies) or volumetric MRI (4 studies), neither technique appears to be consistently better than the other in localising the epileptogenic focus. Very few studies are available which compare PET to CT or MRS.

In studies restricted to patients with temporal lobe epilepsy, heterogeneity is less prominent [230]. In this patient group, ictal SPECT has the best results (correctly localising the epileptogenic site in 70-100% of the scans; and 0-7% of non-localising scans). PET appears to be promising (9 studies), with correct localisation ranging from 56-88% and non-localisation in less than 25%. However, more research is necessary to refine these results.

To evaluate the capacity of PET to predict surgery outcome, five studies were included [230]. None of them showed a statistical significant association between having a correct localisation on PET and having a good surgical outcome. Although a trend was noted towards a better surgical outcome for patients with a correctly localised scan, none of these studies included an appropriate patient spectrum.

In 3 studies a multivariate analysis was performed to investigate the association of PET scan results with the outcome following surgery. This showed the same result.

The authors of this systematic review conclude that these results do little to inform clinical practice, owing to the limitations of the included studies [230].

6.3.3 Primary studies

6.3.3.1 Diagnostic accuracy studies

Of the retrieved primary studies on the diagnostic use of PET for epilepsy surgery, none were ultimately included, mainly because of partial verification (i.e. no verification by a reference standard of patients not undergoing surgery).

Nevertheless, according to the consulted experts, FDG-PET might have a role in detecting surgical candidates among patients with temporal lobe epilepsy but with a negative MRI. PET might also have a role in the clinical follow-up of patients after epilepsy surgery, since it can demonstrate postoperative functional metabolic improvement in regions connected to but remote from the resection area. This metabolic improvement might correlate to cognitive improvement as demonstrated by neuropsychological testing. Likewise, in patients with severe epilepsy and cognitive impairment, it is assumed that better epilepsy control might lead to cognitive improvement and improvement on PET scan in specific regions. This expert opinion is based on several publications that did not fulfil the inclusion criteria of our study [231, 234-236].

Finally, ictal PET has been gaining attention, but this promising technique still needs further scientific evaluation.

6.3.3.2 Prognostic studies

One retrospective study evaluated the prognostic value of PET in 193 consecutive patients with refractory neocortical focal epilepsy undergoing surgery [237]. Multivariate analysis showed that correct localization of the epileptogenic focus by PET (p=0.007) was an independent predictor of a good outcome.

In another prospective prognostic study of moderate quality, Gaillard et al. investigated the prognostic value of PET in 38 children taking anti-epileptic drugs for partial seizures [238]. A normal MRI (OR 0.036, p<0.01 for poor outcome) was found to have a higher predictive value than a normal PET (OR 0.215 p=0.20).

Key messages

- **No new eligible primary studies were identified on the use of PET for the pre-surgical evaluation of refractory epilepsy, but a new systematic review confirms the conclusions of the previous KCE report (level 2). Nevertheless, PET seems to be already embedded in daily practice for this indication (expert opinion).**
- **PET may have a prognostic value in patients undergoing surgery for refractory epilepsy. However, good prospective studies are needed to confirm this. Above this, the clinical consequences are unclear.**
- **In children with epilepsy, PET was not found to be a prognostic factor for seizure control in 1 small study.**

7 PET IN INFECTIOLOGY

The previous KCE report identified only limited evidence on the use of PET in patients with prosthetic hip joint replacement and fever of unknown origin [1]. At that time, evidence was considered too limited to support the widespread use of PET in these indications.

Our search identified evidence on the following indications: osteomyelitis, prosthetic joint infections, fever of unknown origin, infections of the vertebral column and vascular infections.

7.1 OSTEOMYELITIS

The CADTH 2008 report [239] identified 1 systematic review [240] and 3 additional primary studies on the use of PET in patients with suspected osteomyelitis. In the systematic review of Termaat et al., 4 primary studies were included. For PET, pooled sensitivity and specificity were 96% (95%CI 88-99%) and 91% (95%CI 81-95%) respectively for the detection of chronic osteomyelitis [240]. PET was significantly more sensitive than other imaging tests (including leukocyte scintigraphy, bone scintigraphy, and MRI).

The 3 primary studies included in the CADTH 2008 report evaluated PET for the detection of osteomyelitis in different populations and using different comparators [239]. Sensitivity ranged from 29-100%, specificity ranged from 78-92%.

No additional primary studies were identified by our search.

7.2 PROSTHETIC JOINT INFECTIONS

The CADTH 2008 report identified one low-quality systematic review that included 6 primary studies evaluating PET for the detection of prosthetic hip and knee joint infection [239]. PET was found to be the most accurate imaging technique (diagnostic accuracy of 92%), with a sensitivity and specificity of 94% and 87% respectively.

In a recent systematic review, Kwee et al. included 11 primary studies evaluating PET for the detection of prosthetic hip and knee joint infection [241]. Pooled sensitivity and specificity were 82% (95%CI 68-91%) and 87% (95%CI 80-91%) respectively.

No additional primary studies were identified by our search.

7.3 FEVER OF UNKNOWN ORIGIN

The CADTH 2008 report identified one primary study involving 70 patients with fever of unknown origin (FUO) [239]. Sensitivity and specificity for PET were 88% and 77% respectively. PET contributed to the final diagnosis in 33% of patients.

Our search identified one additional prospective study evaluating the use of PET/CT in 48 patients with FUO [242]. Sensitivity and specificity for the detection of an infectious focus were 100% and 81% respectively. This study possibly suffered from incorporation bias.

7.4 INFECTIONS OF THE VERTEBRAL COLUMN

The CADTH 2008 report reported on one low-quality meta-analysis evaluating the use of PET in patients with suspected infection of the vertebral column [239]. Of the investigated diagnostic tests (including CT, MRI, Gallium scan, etc.), PET had the highest sensitivity (100%). Specificity was 88%. However, the results are based on small sample sizes and no confidence intervals are presented.

No additional primary studies were identified by our search.

7.5 VASCULAR INFECTIONS

Our search identified one prospective study evaluating PET/CT for the diagnosis of vascular graft infections [243]. A sensitivity and specificity of 93% and 92% were found. This study possibly suffered from incorporation bias.

Key messages

- **The identified evidence on the use of PET for the diagnosis of osteomyelitis investigated heterogeneous populations. One systematic review suggests a possible role in patients with suspected chronic osteomyelitis, with a sensitivity for PET above 90% (level 2).**
- **For the diagnosis of prosthetic joint infections, results reported in meta-analyses show moderate results for PET (pooled sensitivity 82%, pooled specificity 87%). In view of the absence of better performing imaging techniques and the inability to use MRI in these patients, these results can be considered clinically relevant (level 2).**
- **New evidence on the use of PET in the work-up of patients with fever of unknown origin shows moderate results (level 2). Since this is considered a potential indication for PET(/CT), good prospective studies are needed.**
- **The evidence on the use of PET/CT for the evaluation of prosthetic vascular graft infections is limited to one prospective study showing good results (level 2). Confirmation is needed in additional studies.**

8 CRITERIA FOR PROGRAMMING OF PET

8.1 INTRODUCTION

According to the law of April 27th 2005, the maximum number of nuclear medicine departments in Belgium allowed to have a PET scanner is limited according to pre-specified criteria: 1 department for each university hospital (n = 7), 1 department for each hospital delivering surgical and medical services exclusively in oncology (n = 1), and 1 department per 1.6 million inhabitants (n = 5; 3 departments in the Flemish region and 2 departments in the Walloon region). Taking into account these criteria, the total number of approved PET scanners is 13.

To regulate these 13 PET scanners, specific accreditation criteria are used. These include: a proof of sufficient oncological activity; the presence of a gamma camera; the availability of a medical staff at least including 3 FTE nuclear medicine specialists, 1 FTE physicist or ingeneer, and 2 FTE nurses exclusively active in the nuclear medicine department; internal registration; and external quality control.

As specified in the previous KCE report [1], reimbursement of PET has three components: the annual flat-rate amount paid per approved PET scanner, the medical fee-for-service and the reimbursement of FDG. At present, the medical fee-for-service is only applicable to the 16 indications listed in table 18. These indications have not changed since the previous KCE report.

Table 18. Reimbursed indications for PET according to tumour site or specialism (Source: RIZIV/INAMI).

Lung cancer
Evaluation of solitary pulmonary nodule before surgery
Whole-body examination for initial staging of NSCLC, if the therapy, particularly curative surgery, is decisively influenced by the examination
Whole-body examination for evaluation of residual mass or in case of confirmed suspicion of recurrent NSCLC
Lymphoma
Whole-body examination for initial staging of Hodgkin lymphoma or NHL (intermediary or advanced stage), if the therapy, particularly curative surgery, is decisively influenced by the examination
Whole-body examination for evaluation of residual mass or in case of confirmed suspicion of recurrent lymphoma
Head and neck cancer
Evaluation of residual mass or in case of documented suspicion of recurrent oral or pharyngeal malignancy
Colorectal cancer
Whole-body examination for evaluation of residual mass or in case of confirmed suspicion of recurrent colorectal cancer
Malignant melanoma
Whole-body examination for initial staging of malignant melanoma (stage IIc or above), if the therapy, particularly curative surgery, is decisively influenced by the examination
Whole-body examination for evaluation of residual mass or in case of confirmed suspicion of recurrent malignant melanoma
Oesophageal cancer
Whole-body examination for initial staging of oesophageal cancer, if the therapy, particularly curative surgery, is decisively influenced by the examination
Pancreatic cancer
Whole-body examination for initial staging of pancreatic cancer, if the therapy, particularly curative surgery, is decisively influenced by the examination
Whole-body examination for evaluation of residual mass or in case of confirmed suspicion of recurrent pancreatic cancer
Ovarian cancer
Whole-body examination for evaluation of residual mass or in case of confirmed suspicion of recurrent ovarian cancer

Brain cancer
Evaluation of residual mass or in case of documented suspicion of recurrent cerebral malignancy
Cardiology
Evaluation of myocardial viability if surgery is planned for a recent well-documented coronary insufficiency and if there still is doubt about the myocardial viability
Neurology
Epilepsy that does not respond to medication if the scan can decisively influence the therapeutic management towards curative surgery

Table 19 shows the evolution of the number of reimbursed (i.e. official) PET scans and the associated expenses between 2002 and 2007. A steady grow in expenses can be seen, mainly due to an increase in the number of outpatient PET scans. Importantly, these figures do not represent the real number of PET scans performed in Belgium. In fact, since 2002, 'unofficial' PET scans can be reimbursed through the nomenclature code 'scintigraphy double tomography' (http://www.riziv.fgov.be/care/nl/nomenclature/pdf-IRI/art18IRI.pdf). In the past, this code was created for the reimbursement of certain cardiac scintigraphies. In 2001, 13 721 such double tomographies were performed (i.e. before unofficial PET scan procedures were allowed). In 2007, the number of double tomographies was almost tripled to 34 421 (~€ 9.5 million). Assuming that the number of cardiac double tomographies remained almost stable between 2001 and 2007, about 20 000 double tomographies would be attributable to PET scans in 2007 (~€ 5.5 million), resulting in a total number of PET scans of about 38 500 (official and unofficial). Finally, in 2007 more than 25 000 radio-isotopes with low atomic weight (code 699215 – 699226) were reimbursed, totalling € 4.18 million.

Table 19. Number of reimbursed PET scans and associated expenses (in euro), 2002-2007 (Source: RIZIV/INAMI).

Year	Number of reimbursed PET scans			Expenses
	Outpatient	Inpatient	Total	
2002	7 874	2 357	10 231	1 778 580
2003	9 634	2 822	12 456	1 927 848
2004	10 639	2 779	13 418	2 083 657
2005	12 631	2 962	15 593	2 414 693
2006	13 698	2 871	16 569	2 606 839
2007	15 381	3 107	18 488	2 961 692

Recently, the European Commission received a complaint against the law of April 27th 2005 and the hospital law of August 7th 1987, by which the installation and exploitation of PET scans in departments not fulfilling these criteria (and de facto not being part of the 13 approved PET scans) is prohibited. On January 31st 2008, the European Commission decided to disregard the complaint provided that Belgium adapts its current programming using objective criteria.

In order to allow the development of objective criteria for the programming of PET scan in Belgium, the use of such criteria in other countries was evaluated (chapter 8.1). In addition, possible approaches for the estimation of the number of patients requiring a PET scan are discussed (chapter 8.2).

8.2 USE OF CRITERIA FOR PET PROGRAMMING IN OTHER COUNTRIES

8.2.1 Sources of information

Very recently, INAHTA surveyed their members on the use of standards or guidance for determining how many PET scans are needed in their country (www.inahta.org). We received authorisation to use these results for the present report (Liz Adams, personal communication). In addition, we searched the internet for more recent documents on this issue and contacted experts to provide additional information.

8.2.2 INAHTA survey

On February 13th 2009, INAHTA posted an email survey to all INAHTA members inquiring the use of standards or guidance for determining how many PET scans are needed for their healthcare system (e.g. x numbers of PET scanners / million persons or another method). Fourteen HTA agencies (including the KCE) from 11 different countries responded to this survey (see table 20).

Table 20. Overview of the use of standards/guidance for PET programming in 10 countries (source: INAHTA).

Country	Use of standards/ guidance	Remarks
Argentina	Negative	-
Australia	Yes	At present, Australia uses pre-specified accreditation criteria, and has a list of 9 indications that are reimbursed. This list will be extended with 6 additional indications by May 2009. Two facilities are funded for research.
Canada	Negative	As of November 2007, there were 22 centres performing publicly funded PET scans in seven Canadian provinces. In Ontario, PET scans are currently part of government funded field evaluation studies determining the clinical utility in treatment determinations of PET scan.
Denmark	Negative	-
Germany	Negative	-
Israel	Yes	Israel uses a Certificate of Need (CON) process which calculates the distribution of high cost medical devices (including PET) according to the population size (determined by the Ministry of Health). The need for PET scans is estimated to be 1 per million citizens.
Malaysia	Negative	Is based on the advice of the National Advisor of Nuclear Medicine.
Mexico	Negative	There is 1 PET scan under the organisation of the Ministry of Health (Cancerology National Institute, Mexico City) and 8 PET scans in private hospitals.
Spain	Yes	Each of the 17 Comunidad Autónoma (Regional Governments) has the authority for planning, although there is a certain level of consensus for the indications which are formulated by a national commission and updated in a periodic report. At present, the number of PET scans per million citizens varies between 1 and 3.3 across the 17 regions.
United Kingdom	Negative	In 2005, the Department of Health published a framework for the development of PET services in England. It seeks to provide advice on the current status of PET, the evidence base, number of scanners needed, workforce and training issues, costs and research.

Clearly, only a minority of the countries of which answers were received use criteria for PET programming or have a limited list of reimbursed indications. Israel uses population size to assess the need for PET scans.

In Spain, PET programming is regionalised, although the exact criteria are unclear. However, 'there is a certain level of consensus for the indications'. Australia has a limited list of reimbursed indications (table 21) and uses specified criteria to accredit a facility.

Table 21. Reimbursed PET indications in Australia (source: INAHTA).

Reimbursed PET indication
Evaluation of solitary pulmonary nodule, where the lesion is considered unsuitable for transthoracic fine needle aspiration biopsy, or for which an attempt at pathological characterisation has failed.
Staging of proven NSCLC, where curative surgery or radiotherapy is planned.
Following initial therapy, evaluation of suspected residual, metastatic or recurrent colorectal carcinoma in patients considered suitable for active therapy.
Following initial therapy, evaluation of suspected residual, metastatic or recurrent colorectal carcinoma in patients considered suitable for active therapy, with catheterisation of the bladder.
Following initial therapy, evaluation of suspected metastatic or recurrent malignant melanoma in patients considered suitable for active therapy.
Following initial therapy, evaluation of suspected metastatic or recurrent malignant melanoma in patients considered suitable for active therapy, with catheterisation of the bladder.
Evaluation of refractory epilepsy, which is being evaluated for surgery.
Following initial therapy, evaluation of suspected residual, metastatic or recurrent ovarian carcinoma in patients considered suitable for active therapy.
Following initial therapy, evaluation of suspected residual, metastatic or recurrent ovarian carcinoma in patients considered suitable for active therapy, with catheterisation of the bladder.

8.2.3 The Netherlands

Recently, ZonMw published a report on the use of PET in the Netherlands [2]. By the end of 2006, approximately 24 PET(/CT) scans were available, although not all were operational at that time. Based on the available evidence (until 2006, and assuming that scientific evidence is needed to justify the use of PET for an indication) and incidence data of 2005, ZonMw calculated a need of 17 836 scans or 8.6 PET(/CT) devices [2]. Based on these data, and after discussion with policy makers, professional associations and insurers, ZonMw formulated the following recommendations:

- An efficient use of PET(/CT) should be achieved:
  - Using guidelines and local protocols for routine use of PET(/CT);
  - Using a national registration in order to evaluate the effectiveness and efficiency of PET(/CT) scans outside these guidelines;
  - Performing research for those indications where evidence is lacking.
- The quality of PET(/CT) should be guaranteed:
  - Through continuous education of nuclear medicine specialists, radiologists and prescribers of PET(/CT);
  - Through structured collaboration agreements between nuclear medicine specialists and radiologists.

However, to our knowledge these recommendations are not implemented so far, nor is a list of reimbursed indications being used.

8.3 ESTIMATION OF THE NUMBER OF PATIENTS REQUIRING A PET(/CT) SCAN

8.3.1 Potential population impact

In contrast to the previous KCE report [1], much more precise data are available on the incidence of cancer in Belgium at present. The most recent data date from 2005. Table 22 provides an overview of the incidence of the cancers discussed in this report.

Table 22. Incidence of cancers discussed in this report, 2005 (Source: Belgian Cancer Registry).

Cancer type	ICD-10 code(s)	Incidence
Head-and-neck cancer	C00 – 14, C30 – 32	2 359
Oesophageal cancer	C15	888
Gastric cancer	C16	1 362
Colorectal cancer	C18 – 20	7 519
Primary liver cancer	C22	448
Pancreatic cancer	C25	1 034
Lung cancer	C34, C39, C45	7 062
Malignant melanoma	C43	1 560
Breast cancer	C50	9 486
Cervical cancer	C53	651
Uterine cancer	C54 – 55	1 411
Ovarian cancer	C56	908
Penile cancer	C60	67
Prostate cancer	C61	9 510
Testicular cancer	C62	283
Renal cancer	C64	1 329
Bladder cancer	C67	2 052
Brain cancer	C70 – 71	742
Thyroid cancer	C73	649
Carcinoma of unknown primary	C80	1 175
Lymphoma	C81 – 85	2 125
Hodgkin's disease	*C81*	*295*
Non-Hodgkin-lymphoma	*C82 – 85*	*1 830*

The incidences of the cardiologic, neurologic and infectious indications discussed in this report are much more difficult to estimate. However, the impact for the cardiologic indications is potentially high.

8.3.2 Number of patients requiring a PET(/CT) scan

Broadly, two approaches are possible to estimate the number of patients requiring a PET(/CT) taking into account the scientific evidence.

First, a prospective approach can be chosen, using evidence-based clinical indications and epidemiological data (e.g. cancer incidences, stage distribution, etc.) with future projections. This would be the ideal approach, since the final number would be entirely based on the indications that are based on evidence, no less, no more. For some indications the calculation is fairly straightforward and reliable. For example, PET(/CT) is indicated for the staging of NSCLC. According to the Belgian Cancer Registry data, 9 594 individuals had NSCLC in 2004-2005. This corresponds to 4 800 annual PET scans.

However, Belgian epidemiological data are only systematically available for oncology, and these are often insufficiently detailed. For example, the detection of recurrent disease in patients curatively treated for colorectal cancer is a clear indication for PET(/CT). Curative treatment is mainly limited to patients with pathological stage I-III, totalling 9 637 patients in 2004-2005 (see table 23). However, no exact data exist on how many of these patients had (suspected) recurrent disease during follow-up after curative treatment. This also applies to other similar indications, such as recurrent ovarian cancer and recurrent head-and-neck cancer.

Table 23. Number of new diagnoses of colorectal cancer according to pathological stage, Belgium 2004-2005 (source: Belgian Cancer Registry).

Pathological stage	Men	Women	Total
I	1119	868	1987
II	2119	1834	3953
III	1952	1745	3697
IV	655	544	1199
X	2356	1907	4263

Another indication is the distinction between benign and malignant solitary pulmonary nodules. Where exact data are available on the incidence of lung cancer (NSCLC + SCLC, 13 802 patients in 2004-2005), it is unclear how many of these had a solitary nodule and how many benign solitary nodules needed evaluation with PET scan. The same problem is true for the differential diagnosis between a benign and malignant pancreatic cyst or between chronic pancreatitis and pancreatic cancer.

For malignant melanoma, PET(/CT) is indicated for the M-staging of patients with advanced stages, i.e. stage III and IV. In 2004-2005, 83 patients had clinical stage III or IV malignant melanoma (see table 24). However, these data on clinical stage already take into account the results of PET scans, leading to an underestimation (e.g. a patient with stage III-IV before PET can be downstaged by PET and receive a final clinical stage I-II; this patient would not have been included in the calculations).

Table 24. Number of new diagnoses of malignant melanoma according to clinical stage, Belgium 2004-2005 (source: Belgian Cancer Registry).

Clinical stage	Men	Women	Total
III	23	12	35
IV	30	18	48

A second approach would be to base the calculations on the actual use of PET. This retrospective approach was used in the previous KCE report [1]. An important advantage is that these data are easy to collect, since every department with an approved PET device is obliged to register pre-specified parameters for each performed PET scan, including the indication. The total number of patients requiring a PET scan would be simply calculated by adding up all appropriate PET scans (i.e. PET scans performed for indications based on evidence). However, a major problem is that underuse is neglected with this approach. Indeed, patients requiring but not receiving a PET scan are not present in these registers. Above this, these data are not publicly available. For these reasons, and in contrast to the previous KCE report [1], the data were not used for the present report.

9 DISCUSSION

9.1 METHODOLOGY

During the expert meetings, the methodology of the present report was highly criticised. The essence of this criticism lies in the fact that for another recent KCE report on MRI [244] a health services research approach was chosen, while for the present report a systematic review was done to answer the research question.

Clearly, the difference in approach between the MRI report and the present report on PET scan can be explained by the difference in research question. While the report on MRI examined the issue of programming and financing of MRI, with a focus on the costs associated with running a MRI facility in a Belgian setting [244], the present report mainly answers a clinical question (i.e. what is the diagnostic accuracy and clinical effectiveness of PET and PET/CT). Specifically for PET scan, the present report should be considered as a step-up to a subsequent report focusing on the programming and financing of PET scan. In the present report, a basis is provided for such a subsequent report by providing (in addition to an overview of the clinical indications of PET and PET/CT) an overview of the programming criteria used in other countries and by discussing the caveats when calculating the number of patients requiring a PET scan.

Another point of criticism was the use of stringent criteria to select the evidence. However, it is clear that such criteria are necessary to reduce possible bias. It is known that shortcomings in study design (e.g. non-consecutive inclusion of patients, retrospective data collection, patient selection based on referral for index test, etc.) can affect (i.e. increase) estimates of diagnostic accuracy [16]. However, we acknowledge the fact that some sources of bias cannot be eliminated, such as differential verification (i.e. the use of two or more different reference standards, e.g. histology or clinical follow-up). Some of these were therefore not used as an exclusion criterion.

Importantly, it is a misunderstanding that we limited our search to synthesized evidence and RCTs. Questions about the diagnostic accuracy of diagnostic tests can be answered with good observational studies. In fact, most indications discussed in the present report and for which PET(/CT) is considered appropriate, have an evidence base that is limited to observational studies.

Since it was not our intention to develop clinical practice guidelines on PET and PET/CT but rather to focus on the conclusions of the literature on PET and PET/CT, we did not systematically search for guidelines. Rather, the present report should be considered as a basis to develop national guidelines. Nevertheless, we agree that the approach in the present report was restricted by not systematically taking into account other diagnostic techniques. However, this was due to the specific search question and time constraints. Indeed, a broader approach would imply to do a systematic review for all relevant diagnostic techniques, such as CT and MRI. This was impossible within this short timeframe.

This restricted approach lead to some conclusions that were considered to be too negative by the consulted experts. If, for example, a sensitivity or specificity of 60% was found for PET, this was considered to be low as such by the authors. However, it is of course possible that other imaging techniques have even lower sensitivities or specificities. In that case PET would be the preferred technique, even with such a low diagnostic accuracy. However, this also applies in the opposite direction. If PET was found to have a sensitivity or specificity of 85%, this was considered to be moderate as such by the authors. If other techniques would have better diagnostic accuracy, PET would not be recommended. As a consequence, ‘moderate’ would have been too positive in that case.

9.2 CLINICAL INDICATIONS FOR PET AND PET/CT

Currently, the reimbursement of PET scan in Belgium is limited to 16 indications. Ten of these indications are supported by evidence. However, for some indications, the literature is too limited to support the use of PET at present (table 25).

Table 25. Reimbursed indications for PET(/CT) and their support by evidence.

Reimbursed indication	Evidence (level) #
Lung cancer	
Evaluation of solitary pulmonary nodule	Yes (level 2)
Initial staging of NSCLC	Yes (level 4)
Evaluation of residual mass or recurrent NSCLC	Inconclusive (level 2)
Lymphoma	
Initial staging of Hodgkin lymphoma or NHL (intermediary or advanced stage)	Yes (level 2)
Evaluation of residual mass or recurrent lymphoma	Yes for residual mass (level 2) Inconclusive for recurrent lymphoma (level 2)
Head and neck cancer	
Evaluation of residual mass or recurrent oral or pharyngeal malignancy	Yes (level 2)
Colorectal cancer	
Evaluation of residual mass or recurrent colorectal cancer	Yes (level 3)
Malignant melanoma	
Initial staging of malignant melanoma (stage IIc or above)	Yes (level 2)
Evaluation of residual mass or recurrent malignant melanoma	Inconclusive for residual mass and detection of recurrence (level 2) Yes for staging of recurrence (level 2)
Oesophageal cancer	
Initial staging of oesophageal cancer	Yes (level 2)
Pancreatic cancer	
Initial staging of pancreatic cancer	Inconclusive (level 2)
Evaluation of residual mass or recurrent pancreatic cancer	No (level 2)
Ovarian cancer	
Evaluation of residual mass or recurrent ovarian cancer	Yes (level 2)
Brain cancer	
Evaluation of residual mass or recurrent cerebral malignancy	Inconclusive (level 2)
Cardiology	
Evaluation of myocardial viability	Yes (level 2)
Neurology	
Preoperative evaluation of refractory epilepsy	Inconclusive (level 2)

As discussed in the Methodology section, 4 levels of diagnostic accuracy are distinguished: (1) technical accuracy; (2) diagnostic accuracy; (3) impact on patient outcome; (4) cost-effectiveness.

On the other hand, new indications for PET and PET/CT have emerged in comparison to the current list of reimbursed indications (table 26). In addition, some indications are potentially relevant and warrant further research (table 27). Worth mentioning here is the potential use of PET and PET/CT in the monitoring of treatment response, as suggested by the external experts. Finally, at present PET and PET/CT are not indicated for primary liver cancer, gastric cancer, breast cancer, testicular cancer, bladder cancer, prostate cancer, uterine cancer and penile cancer.

This evolution in indications and the bulk of new evidence published in the last 4 years underpin the need to regularly update the evidence on PET and PET/CT. Ideally, as stated above, such an update should position PET and PET/CT against other diagnostic techniques in a systematic way. Again, this approach was impossible for the present report because of the specific search question (focused on PET and PET/CT) and time constraints.

Table 26. New indications for PET(/CT) supported by conclusive level 2 evidence, according to tumour site.

Head and neck cancer
Primary staging
Colorectal cancer
Preoperative evaluation of potentially resectable colorectal liver metastases
Pancreatic cancer
Differentiation between chronic pancreatitis and pancreatic cancer and between benign and malignant pancreatic cysts
Cervical cancer
Primary staging
GIST
Treatment monitoring
Carcinoma of unknown primary
Detection of primary tumour

Table 27. Potential indications for PET(/CT) (supported by inconclusive level 2 evidence), according to tumour site or specialism.

Lung cancer
Radiotherapy planning
Oesophageal cancer
Monitoring of treatment response
Thyroid cancer
Evaluation of thyroid nodules with inconclusive cytological results
Cervical cancer
Evaluation of recurrent cervical cancer
Ovarian cancer
Initial diagnosis
Infectious diseases
Evaluation of chronic osteomyelitis
Evaluation of prosthetic joint infections
Evaluation of fever of unknown origin

Overall, the evidence on PET and PET/CT is mainly limited to diagnostic accuracy studies. Some of these studies report a change in management, although it is not clear how these management changes affect patient outcomes. That is also the reason why the levels of evidence used in the previous KCE report [1] were not used in the present report (see chapter 3.5). Indeed, to evaluate the effect of changes in management on patient outcomes, the most appropriate design would be a RCT. However, only 2 new RCTs were identified, one on the detection of recurrent colorectal cancer [99] and one on the evaluation of myocardial viability [214].

Although the evidence for new indications has increased during the last 4 years, no clear positive evolution was seen in the quality of the evidence. Many of the included studies suffered from methodological flaws. Most studies had relatively small sample sizes. Studies were also often retrospective, and some of them suffered from incorporation bias. For researchers this should be considered as an important trigger to improve the quality of the evidence in this domain.

9.3 PROGRAMMING OF PET

Belgium is one of the rare countries that use a programming policy for PET scanners. Some countries, such as Israel and France (see previous KCE report [1]), limit the number of PET scanners based on population size. Accreditation criteria are used by some countries, such as Australia and France (see previous KCE report [1]), to regulate PET scan. Finally, some countries, such as Australia, Spain and the USA (see previous KCE report [1]), have a limited list of reimbursed indications. To our knowledge, Belgium is the only country that combines all these criteria and modalities, and in addition programs the maximum number of PET scanners based on pre-defined criteria.

The evolution of the number of reimbursed and unofficial PET scans clearly shows that the present programming of PET scanners did not result in a controlled increase in the number of examinations. On the contrary, the number of unofficial PET scans exploded since 2002, and was at least as high as the number of reimbursed PET scans in 2007.

Basing the programming of PET on the number of patients requiring a PET scan is difficult, since no reliable data exist to calculate this number and since the evidence-based indications are in constant evolution. For some indications, this calculation is fairly straightforward, but for other indications (e.g. evaluation of recurrent disease) it is impossible. In The Netherlands, ZonMw estimated the number of patients requiring a PET scan to be 17 836 for the year 2005 [2]. This estimation was mainly based on several separate projects financed by ZonMw. Due to time constraints, this approach was impossible for the present study.

9.4 CONCLUSIONS AND POLICY RECOMMENDATIONS

Four important conclusions can be drawn from the present report:

1. During the last 4 years, the body of evidence for new indications, currently not reimbursed in Belgium, has increased although the quality of this evidence did not improve.
2. By allowing the reimbursement of PET scans through the nomenclature code 'scintigraphy double tomography', the programming of PET scanners had only a minor influence on the real number of PET examinations.
3. Two methods are available to align the number of PET scanners to the clinical needs, i.e. programming on the one hand and accreditation criteria and reimbursement modalities on the other hand.
4. In Belgium, a programming of PET scanners based on a calculation of the needs is impossible in the short run.

Policy recommendations

- Calculating the number of PET scans needed is impossible in the short run and is therefore not recommended as a means of programming PET scanners in Belgium.
- An alternative to programming is to regulate the number of PET scanners:
    - by setting accreditation criteria that are strict enough and of which the application is strictly monitored to assure the quality of the examinations;
    - by determining reimbursement criteria that limit the reimbursable indications to those that are based on scientific evidence.
- The reimbursement of PET examinations is conditional upon registration of the indication in a unique, authomised and standardised registry. This mandatory registration should allow to follow up if the proposed system corresponds to the actual needs.
- The limitative list of reimbursed indications for PET and PET/CT should be updated every three years, with special attention for new tracers and new imaging modalities. For this 3-yearly update, the research question should be expanded to other imaging techniques, to allow a systematic positioning of PET and PET/CT towards these techniques.
- If an oncological indication supported by inconclusive scientific evidence is added to this list, reimbursement should be linked to the multidisciplinary oncological consult.
- The compliance with these reimbursement criteria should be checked systematically.
- Reimbursement of PET scans through the nomenclature code 'scintigraphy double tomography' should be abandoned to allow a transparent and controlled follow-up of the number of PET investigations.

10 APPENDICES

APPENDIX 1: SEARCH STRATEGIES

ADAPTED MIJNHOUT STRATEGY

1	deoxyglucose/ or deoxyglucose.tw. or desoxyglucose.tw. or deoxy-glucose.tw. or desoxy-glucose.tw. or deoxy-d-glucose.tw. or desoxy-d-glucose.tw. or 2deoxyglucose.tw. or 2deoxy-d-glucose.tw. or fluorodeoxyglucose.tw. or fluorodesoxyglucose.tw. or fludeoxyglucose.tw. or fluordeoxyglucose.tw. or fluordesoxyglucose.tw. or 18fluorodeoxyglucose.tw. or 18fluorodesoxyglucose.tw. or 18fluordeoxyglucose.tw. or fdg*.tw. or 18fdg*.tw. or 18f-dg*.tw.
2	Fluorodeoxyglucose F18/
3	(fluor or 2fluor* or fluoro or fluorodeoxy or fludeoxy or fluorine or 18f or 18flu*).tw.
4	glucose.tw.
5	3 and 4
6	1 or 2 or 5
7	Positron-Emission Tomography/
8	(pet or petscan*).tw. or tomography, emission-computed/
9	emission.tw.
10	(tomograph or tomographs or tomographic* or tomography or tomographies).tw.
11	9 and 10
12	7 or 8 or 11
13	6 and 12
14	animals/ not humans/
15	13 not 14

GENERIC SEARCH ONCOLOGY

1	exp Neoplasms/
2	Neoplasm Staging/
3	cancer$.ti,ab.
4	tumor$.ti,ab.
5	tumour$.ti,ab.
6	carcinoma$.ti,ab.
7	neoplasm$.ti,ab.
8	lymphoma.ti,ab.
9	melanoma.ti,ab.
10	staging.ti,ab.
11	metastas$.ti,ab.
12	metastatic.ti,ab.
13	exp Neoplasm Metastasis/
14	exp neoplastic processes/
15	neoplastic process$.ti,ab.
16	non small cell.ti,ab.
17	adenocarcinoma$.ti,ab.
18	squamous cell.ti,ab.
19	nsclc.ti,ab.
20	osteosarcoma$.ti,ab.
21	phyllodes.ti,ab.
22	cystosarcoma$.ti,ab.
23	fibroadenoma$.ti,ab.
24	(non adj small adj cell).ti,ab.
25	(non adj2 small adj2 cell).ti,ab.
26	(nonsmall adj2 cell).ti,ab.
27	plasmacytoma$.ti,ab.
28	myeloma.ti,ab.
29	multiple myeloma.ti,ab.

30	lymphoblastoma$.ti,ab.
31	lymphocytoma$.ti,ab.
32	lymphosarcoma$.ti,ab.
33	immunocytoma.ti,ab.
34	sarcoma$.ti,ab.
35	hodgkin$.ti,ab.
36	(nonhodgkin$ or non hodgkin$).ti,ab.
37	or/1-36

SYSTEMATIC REVIEWS AND META-ANALYSES: OVID MEDLINE

1	Meta-Analysis as Topic/
2	meta analy$.tw.
3	metaanaly$.tw.
4	Meta-Analysis/
5	(systematic adj (review$1 or overview$1)).tw.
6	exp Review Literature as Topic/
7	or/1-6
8	cochrane.ab.
9	embase.ab.
10	(psychlit or psyclit).ab.
11	(psychinfo or psycinfo).ab.
12	(cinahl or cinhal).ab.
13	science citation index.ab.
14	bids.ab.
15	cancerlit.ab.
16	or/8-15
17	reference list$.ab.
18	bibliograph$.ab.
19	hand-search$.ab.
20	relevant journals.ab.
21	manual search$.ab.
22	or/17-21
23	selection criteria.ab.
24	data extraction.ab.
25	23 or 24
26	Review/
27	25 and 26
28	Comment/
29	Letter/
30	Editorial/
31	animal/
32	human/
33	31 and 32
34	31 not 33
35	or/28-30,34
36	22 or 27 or 16 or 7
37	36 not 35

SYSTEMATIC REVIEWS AND META-ANALYSES: EMBASE

(((('computer assisted emission tomography'/exp OR 'positron emission tomography'/exp OR 'whole body tomography'/exp) OR 'positron emission tomography':ti,ab,tn,mn,de OR pet*:ti,ab,de,tn,mn OR petscan*:ti,ab,de,tn,mn OR (pet*:ti,ab,tn,mn,de NOT (animal:ti,ab,tn,mn,de NOT human:ti,ab,tn,mn,de AND animal:ti,ab,tn,mn,de))) AND (('deoxyglucose':ti,ab,de,mn,tn OR 'desoxyglucose':ti,ab,de,mn,tn OR 'deoxy-glucose':ti,ab,de,mn,tn OR 'desoxy-glucose':ti,ab,de,mn,tn OR 'deoxy-d-glucose':ti,ab,de,mn,tn OR 'desoxy-d-glucose':ti,ab,de,mn,tn OR '2deoxyglucose':ti,ab,de,mn,tn OR '2deoxy-d-

glucose':ti,ab,de,mn,tn OR 'fluorodeoxyglucose':ti,ab,de,mn,tn OR 'fluorodesoxyglucose':ti,ab,de,mn,tn OR 'fludeoxyglucose':ti,ab,de,mn,tn OR 'fluordeoxyglucose':ti,ab,de,mn,tn OR 'fluordesoxyglucose':ti,ab,de,mn,tn OR '18fluorodeoxyglucose':ti,ab,de,mn,tn OR '18fluorodesoxyglucose':ti,ab,de,mn,tn OR '18fluordeoxyglucose':ti,ab,de,mn,tn OR fdg*:ti,ab,de,mn,tn OR 18fdg*:ti,ab,de,mn,tn OR '18fdg':ti,ab,de,mn,tn OR ((fluor:ti,ab,de,mn,tn OR 2fluor*:ti,ab,de,mn,tn OR fluoro:ti,ab,de,mn,tn OR fluorodeoxy:ti,ab,de,mn,tn OR fludeoxy:ti,ab,de,mn,tn OR fluorine:ti,ab,de,mn,tn OR 18f:ti,ab,de,mn,tn OR 18flu*:ti,ab,de,mn,tn) AND glucose:ti,ab,de,mn,tn)) OR ('deoxyglucose'/exp OR 'deoxyglucose'))) AND ([meta analysis]/lim OR [systematic review]/lim) AND ([dutch]/lim OR [english]/lim OR [french]/lim) AND [embase]/lim AND [2005-2009]/py

RANDOMISED CONTROLLED TRIALS

1	Randomized controlled trials/
2	Randomized controlled trial.pt.
3	Random allocation/
4	Double blind method/
5	Single blind method/
6	Clinical trial.pt.
7	exp clinical trials/
8	or/1-7
9	(clinic$ adj trial$1).tw.
10	((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw.
11	Placebos/
12	Placebo$.tw.
13	Randomly allocated.tw.
14	(allocated adj2 random).tw.
15	or/9-14
16	8 or 15
17	Case report.tw.
18	Letter.pt.
19	Historical article.pt.
20	Review of reported cases.pt.
21	Review, multicase.pt.
22	or/17-21
23	16 not 22

DIAGNOSTIC ACCURACY STUDIES

1	exp "Sensitivity and Specificity"/
2	sensitivity.tw.
3	specificity.tw.
4	((pre-test or pretest) adj probability).tw.
5	post-test probability.tw.
6	predictive value$.tw.
7	likelihood ratio$.tw.
8	or/1-7

PROGNOSTIC STUDIES

1	(prognos$ or outcome$ or follow-up or predict$).ti,ab,sh.
2	exp disease progression/
3	((natural$ or disease$) adj (progress$ or course$ or histor$)).ti,ab,sh.
4	or/1-3
5	exp cohort studies/
6	(cohort$ or compar$ or longitudinal$ or prospective$ or multivariate or reproducib$).ti,ab,sh.
7	6 or 5
8	4 and 7

SPECIFIC SEARCHES

Breast cancer

Search date: February 12th 2009

1	breast/ or breast diseases/
2	Neoplasms/
3	1 and 2
4	exp Breast Neoplasms/
5	(breast$ adj5 neoplas$).tw.
6	(breast$ adj5 cancer$).tw.
7	(breast$ adj5 carcin$).tw.
8	(breast$ adj5 tumo$).tw.
9	(breast$ adj5 metasta$).tw.
10	(breast$ adj5 malig$).tw.
11	exp Carcinoma, Ductal, Breast/
12	or/3-11

Colorectal cancer

Search date: February 12th 2009

1	(colorectal adj5 neoplas$).tw.
2	(colorectal adj5 cancer$).tw.
3	(colorectal adj5 carcin$).tw.
4	(colorectal adj5 tumo$).tw.
5	(colorectal adj5 metasta$).tw.
6	(colorectal adj5 malig$).tw.
7	(colon$ adj5 neoplas$).tw.
8	(colon$ adj5 cancer$).tw.
9	(colon$ adj5 carcin$).tw.
10	(colon$ adj5 tumo$).tw.
11	(colon$ adj5 metasta$).tw.
12	(colon$ adj5 malig$).tw.
13	(rect$ adj5 neoplas$).tw.
14	(rect$ adj5 cancer$).tw.
15	(rect$ adj5 carcin$).tw.
16	(rect$ adj5 tumo$).tw.
17	(rect$ adj5 metasta$).tw.
18	(rect$ adj5 malig$).tw.
19	(intestin$ adj5 neoplas$).tw.
20	(intestin$ adj5 cancer$).tw.
21	(intestin$ adj5 carcin$).tw.
22	(intestin$ adj5 tumo$).tw.
23	(intestin$ adj5 metasta$).tw.
24	(intestin$ adj5 malig$).tw.
25	(bowel adj5 neoplas$).tw.
26	(bowel adj5 cancer$).tw.
27	(bowel adj5 carcin$).tw.
28	(bowel adj5 tumo$).tw.
29	(bowel adj5 metasta$).tw.
30	(bowel adj5 malig$).tw.
31	exp Rectal Neoplasms/
32	exp Colorectal Neoplasms/
33	exp Colonic Neoplasms/
34	exp Sigmoid Neoplasms/
35	or/22-55

Lung cancer

Search date: February 25th 2009

1	(lung adj5 neoplas$).tw.
2	(lung adj5 cancer$).tw.
3	(lung adj5 carcin$).tw.
4	(lung adj5 tumo$).tw.
5	(lung adj5 metasta$).tw.
6	(lung adj5 malig$).tw.
7	exp Carcinoma, Non-Small-Cell Lung/ or exp Lung Neoplasms/ or exp Small Cell Lung Carcinoma/
8	exp Mesothelioma/
9	exp Bronchial Neoplasms/
10	exp Pleural Neoplasms/
11	NSCLC.tw.
12	SCLC.tw.
13	or/1-12

Lymphoma

Search date: February 25th 2009

1	(hematol$ adj5 neoplas$).tw.
2	(hematol$ adj5 cancer$).tw.
3	(hematol$ adj5 tumo$).tw.
4	(hematol$ adj5 metasta$).tw.
5	(hematol$ adj5 malig$).tw.
6	exp Lymphoma, Non-Hodgkin/ or exp Hodgkin Disease/
7	exp Lymphoma, Mantle-Cell/ or exp Lymphoma, Follicular/ or exp Lymphoma/ or exp Lymphoma, T-Cell, Peripheral/ or exp Leukemia-Lymphoma, Adult T-Cell/ or exp Precursor B-Cell Lymphoblastic Leukemia-Lymphoma/ or exp Burkitt Lymphoma/ or exp Lymphoma, T-Cell, Cutaneous/ or exp Lymphoma, Primary Cutaneous Anaplastic Large Cell/ or exp Lymphoma, Extranodal NK-T-Cell/ or exp Lymphoma, AIDS-Related/ or exp Lymphoma, Large-Cell, Anaplastic/ or exp Precursor T-Cell Lymphoblastic Leukemia-Lymphoma/ or exp Lymphoma, T-Cell/ or exp Lymphoma, Primary Effusion/ or exp Lymphoma, Large B-Cell, Diffuse/ or exp Lymphoma, B-Cell/ or exp Lymphoma, Large-Cell, Immunoblastic/ or exp Precursor Cell Lymphoblastic Leukemia-Lymphoma/ or exp Lymphoma, B-Cell, Marginal Zone/
8	(haematol$ adj5 cancer$).tw.
9	(haematol$ adj5 neoplas$).tw.
10	(haematol$ adj5 tumo$).tw.
11	(haematol$ adj5 metasta$).tw.
12	(haematol$ adj5 malig$).tw.
13	NHL.tw.
14	or/1-13

Pancreatic cancer

Search date: February 25th 2009

1	exp Pancreatic Neoplasms/
2	(pancrea$ adj5 neoplas$).tw.
3	(pancrea$ adj5 cancer$).tw.
4	(pancrea$ adj5 carcin$).tw.
5	(pancrea$ adj5 tumo$).tw.
6	(pancrea$ adj5 metasta$).tw.
7	(pancrea$ adj5 malig$).tw.
8	or/1-7

Oesophageal cancer

Search date: February 25th 2009

1	exp esophageal neoplasms/
2	(esophag$ adj5 neoplas$).tw.
3	(oesophag$ adj5 neoplas$).tw.
4	(esophag$ adj5 cancer$).tw.
5	(oesophag$ adj5 cancer$).tw.
6	(esophag$ adj5 carcin$).tw.
7	(oesophag$ adj5 carcin$).tw.
8	(esophag$ adj5 tumo$).tw.
9	(oesophag$ adj5 tumo$).tw.
10	(esophag$ adj5 metasta$).tw.
11	(oesophag$ adj5 metasta$).tw.
12	(esophag$ adj5 malig$).tw.
13	(oesophag$ adj5 malig$).tw.
14	or/1-13

Thyroid cancer

Search date: February 26th 2009

1	(thyroid$ adj5 neoplas$).tw.
2	(thyroid$ adj5 cancer$).tw.
3	(thyroid$ adj5 carcin$).tw.
4	(thyroid$ adj5 tumo$).tw.
5	(thyroid$ adj5 metasta$).tw.
6	(thyroid$ adj5 malig$).tw.
7	exp Thyroid Neoplasms/
8	or/1-7

Prostate cancer

Search date: March 19th 2009

1	exp Prostatic Neoplasms/
2	(prostate adj5 neoplas$).tw.
3	(prostate adj5 cancer$).tw.
4	(prostate adj5 carcin$).tw.
5	(prostate adj5 tumo$).tw.
6	(prostate adj5 metasta$).tw.
7	(prostate adj5 malig$).tw.
8	or/1-7

Bladder cancer

Search date: March 19th 2009

1	exp Urinary Bladder Neoplasms/
2	exp Carcinoma, Transitional Cell/
3	(bladder adj5 neoplas$).tw.
4	(bladder adj5 cancer$).tw.
5	(bladder adj5 carcin$).tw.
6	(bladder adj5 tumo$).tw.
7	(bladder adj5 metasta$).tw.
8	(bladder adj5 malig$).tw.
9	or/1-8

Gastric cancer

Search date: March 19th 2009

1	exp Stomach Neoplasms/
2	(gastric adj5 neoplas$).tw.
3	(stomach adj5 neoplas$).tw.
4	(gastric adj5 cancer$).tw.
5	(stomach adj5 cancer$).tw.
6	(gastric adj5 carcin$).tw.
7	(stomach adj5 carcin$).tw.
8	(gastric adj5 tumo$).tw.
9	(stomach adj5 tumo$).tw.
10	(gastric adj5 metasta$).tw.
11	(stomach adj5 metasta$).tw.
12	(gastric adj5 malig$).tw.
13	(stomach adj5 malig$).tw.
14	or/1-13

Uterine cancer

Search date: March 19th 2009

1	(uter$ adj5 neoplas$).tw.
2	(endometri$ adj5 neoplas$).tw.
3	(uter$ adj5 cancer$).tw.
4	(endometri$ adj5 cancer$).tw.
5	(uter$ adj5 carcin$).tw.
6	(endometri$ adj5 carcin$).tw.
7	(uter$ adj5 tumo$).tw.
8	(endometri$ adj5 tumo$).tw.
9	(uter$ adj5 metasta$).tw.
10	(endometri$ adj5 metasta$).tw.
11	(uter$ adj5 malig$).tw.
12	(endometri$ adj5 malig$).tw.
13	exp Uterine Neoplasms/
14	exp Endometrial Neoplasms/
15	or/1-14

Epilepsy

Search date: March 19th 2009

1	exp Epilepsy, Partial, Motor/ or exp Epilepsy, Temporal Lobe/ or exp Epilepsy, Complex Partial/ or exp Epilepsy, Partial, Sensory/ or exp Epilepsy, Reflex/ or exp Epilepsy, Benign Neonatal/ or exp Epilepsy, Tonic-Clonic/ or exp Epilepsy, Post-Traumatic/ or exp Epilepsy, Absence/ or exp Epilepsy, Frontal Lobe/ or exp Epilepsy/ or exp Epilepsy, Rolandic/ or exp Myoclonic Epilepsy, Juvenile/ or exp Epilepsy, Generalized/

Malignant melanoma

Search date: March 20th 2009

1	(skin adj5 neoplas$).tw.
2	(skin adj5 cancer$).tw.
3	(skin adj5 carcin$).tw.
4	(skin adj5 tumo$).tw.
5	(skin adj5 metasta$).tw.
6	(skin adj5 malig$).tw.
7	exp Melanoma/
8	exp Skin Neoplasms/
9	or/1-8

Cervical cancer

Search date: March 20th 2009

1	(cervi$ adj5 neoplas$).tw.
2	(cervi$ adj5 cancer$).tw.
3	(cervi$ adj5 carcin$).tw.
4	(cervi$ adj5 tumo$).tw.
5	(cervi$ adj5 metasta$).tw.
6	(cervi$ adj5 malig$).tw.
7	exp Uterine Cervical Neoplasms/
8	or/1-7

Ovarian cancer

Search date: March 20th 2009

1	(ovar$ adj5 neoplas$).tw.
2	(ovar$ adj5 cancer$).tw.
3	(ovar$ adj5 carcin$).tw.
4	(ovar$ adj5 tumo$).tw.
5	(ovar$ adj5 metasta$).tw.
6	(ovar$ adj5 malig$).tw.
7	exp Ovarian Neoplasms/
8	or/1-7

Renal cancer

Search date: March 20th 2009

1	(renal adj5 neoplas$).tw.
2	(kidney adj5 neoplas$).tw.
3	(renal adj5 cancer$).tw.
4	(kidney adj5 cancer$).tw.
5	(renal adj5 carcin$).tw.
6	(kidney adj5 carcin$).tw.
7	(renal adj5 tumo$).tw.
8	(kidney adj5 tumo$).tw.
9	(renal adj5 metasta$).tw.
10	(kidney adj5 metasta$).tw.
11	(renal adj5 malig$).tw.
12	(kidney adj5 malig$).tw.
13	exp Kidney Neoplasms/
14	or/1-13

Parkinson's disease

Search date: March 27th 2009

1	exp Parkinson Disease/

Cardiology

Search date: March 30th 2009

1	exp Myocardial Ischemia/
2	exp Heart Diseases/
3	((heart or coronary or myocardial or cardiac or left ventricular) and (perfusion or viability or metabolism)).mp.
4	or/1-3

Infectious diseases

Search date: March 30th 2009

1	exp Osteomyelitis/
2	exp Prosthesis-Related Infections/
3	exp Joint Prosthesis/ae [Adverse Effects]
4	exp Knee Prosthesis/ae [Adverse Effects]
5	exp Hip Prosthesis/ae [Adverse Effects]
6	exp "Fever of Unknown Origin"/
7	exp Thoracic Vertebrae/ or exp Lumbar Vertebrae/ or exp Spinal Diseases/
8	or/1-7

GIST

Search date: April 2nd 2009

1	exp Gastrointestinal Stromal Tumors/ or GIST.mp.

Primary liver cancer

Search date: April 2nd 2009

1	(liver adj5 neoplas$).tw.
2	(liver adj5 cancer$).tw.
3	(liver adj5 carcin$).tw.
4	(liver adj5 tumo$).tw.
5	(liver adj5 metasta$).tw.
6	(liver adj5 malig$).tw.
7	exp Liver Neoplasms/
8	exp Carcinoma, Hepatocellular/
9	or/1-8

Dementia

Search date: May 7th 2009

1	exp Alzheimer Disease/
2	exp Delirium, Dementia, Amnestic, Cognitive Disorders/ or exp Dementia, Vascular/ or exp Dementia, Multi-Infarct/ or exp Dementia/
3	or/1-2

Brain cancer

Search date: May 19th 2009

1	(brain adj5 neoplas$).tw.
2	(brain adj5 cancer$).tw.
3	(brain adj5 carcin$).tw.
4	(brain adj5 tumo$).tw.
5	(brain adj5 metasta$).tw.
6	(brain adj5 malig$).tw.
7	exp Brain Neoplasms/
8	or/1-7

APPENDIX 2: QUALITY APPRAISAL

HTA REPORTS

INAHTA checklist	HAS 2006	NCCHTA 2007	AHRQ 2008	MSAC 2008	MAS 2005	CADTH 2008	NHS 2006	AHTAPol 2006
Appropriate contact details for further information?	Yes	Yes	Partly	Yes	Yes	Yes	Yes	Yes
Authors identified?	Yes	Yes	Yes	Yes	No	Yes	Yes	No
Statement regarding conflict of interest?	No	Yes	Yes	No	No	No	Yes	Yes
Statement on whether report externally reviewed?	Partly	Yes	No	No	No	Yes	Partly	No
Short summary in non-technical language?	Partly	Yes	Yes	Yes	Yes	Yes	No	Yes
Reference to the policy question that is addressed?	Yes	Partly	Partly	Yes	Yes	Yes	Yes	No
Reference to the research question that is addressed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Scope of the assessment specified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Description of the assessed health technology?	Yes	Yes	No	Yes	Yes	Yes	Partly	Yes
Details on source of information and literature search strategies provided?	Partly 9/10	Partly 8/10	Partly 7/10	Partly 9/10	Partly 7/10	Partly 4/10	Yes 10/10	Partly 7/10
Information on basis for the assessment and interpretation of selected data information?	Partly 2/4	Yes 4/4	Yes 4/4	Yes 4/4	Yes 4/4	Partly 2/4	Yes 4/4	Partly 3/4
Information on context	Partly 4/5	Partly 2/5	No	Partly 2/5	Partly 2/5	No	Partly 1/5	Partly 1/5
Findings of the assessment discussed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Conclusions from assessment clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partly
Suggestions for further action?	Yes	Partly	Yes	Yes	Partly	Yes	Yes	No
Overall appraisal	**Good ***	**Good**	**Good**	**Good**	**Fair**	**Fair**	**Fair**	**Poor**

* Not discussed in the text, because same evidence as discussed in other HTA reports and SR.

SYSTEMATIC REVIEWS

Lung cancer

Dutch Cochrane checklist	Wahidi 2007	Cronin 2008	Berghmans 2008	De Geus 2007	Samson 2005	Ung 2007	Schimmer 2006
Adequate search question?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adequate search strategy?	Yes	Yes	Partly	Partly	Yes	Yes	Yes
Adequate study selection?	Yes	Yes	Yes	Yes	Yes	Yes	No
Adequate quality appraisal?	Yes	Yes	Yes	No	Yes	Yes	No
Adequate description of data extraction?	Yes	Yes	Yes	NA	Yes	NA	NA
Description of most important features of included studies?	Yes	Yes	Yes	No	Yes	Yes	Yes
Correct meta-analysis?	Yes	Yes	Yes	NA	Yes	NA	NA
Valid and applicable systematic review?	**Yes**	**Yes**	**Yes**	**No**	**Yes**	**Yes**	**Yes**

Lymphoma

Dutch Cochrane checklist	Kwee 2008	Pakos 2005	Zijlstra 2006	Terasawa 2008	Brepoels 2007
Adequate search question?	Yes	Yes	Yes	Yes	Yes
Adequate search strategy?	Yes	Yes	Yes	Yes	No
Adequate study selection?	Yes	Yes	Yes	Yes	Yes
Adequate quality appraisal?	Yes	Yes	Yes	Yes	Yes
Adequate description of data extraction?	NA	Yes	Yes	Yes	Yes
Description of most important features of included studies?	Yes	Yes	Yes	Yes	Yes
Correct meta-analysis?	NA	Yes	Yes	Yes	Yes
Valid and applicable systematic review?	**Yes**	**Yes**	**Yes**	**Yes**	**No**

Head & neck cancer

Dutch Cochrane checklist	Kyzas 2008	Brouwer 2008	Isles 2008	Liu 2007	Delgado Bolton 2006
Adequate search question?	Yes	Yes	Yes	Yes	Yes
Adequate search strategy?	Yes	Yes	Yes	Yes	Partly
Adequate study selection?	Yes	Yes	Yes	Yes	Yes
Adequate quality appraisal?	Yes	Yes	Yes	Yes	Yes
Adequate description of data extraction?	Yes	Yes	Yes	Yes	Yes
Description of most important features of included studies?	Yes	Yes	Yes	No	Yes
Correct meta-analysis?	Yes	Yes	Yes	Yes	No
Valid and applicable systematic review?	**Yes**	**Yes**	**Yes**	**Yes**	**No**

Colorectal cancer

Dutch Cochrane checklist	Fletcher 2008	Zhang 2008	Bipat 2005	Wiering 2005	Bipat 2007
Adequate search question?	Yes	Yes	Yes	Yes	Yes
Adequate search strategy?	Yes	Yes	Yes	Yes	Yes
Adequate study selection?	Partly	Yes	Yes	Yes	NA
Adequate quality appraisal?	Partly	Yes	Yes	Yes	NA
Adequate description of data extraction?	Yes	Yes	Yes	No	No
Description of most important features of included studies?	Partly	Yes	No	Yes	No
Correct meta-analysis?	NA	Yes	Yes	No	No
Valid and applicable systematic review?	**Partly ***	**Yes**	**Yes**	**Yes**	**No**

* Not discussed in the text, because just descriptive, only including synthesized evidence and same evidence as other included HTA reports and SR.

Breast and gynaecological cancer

Dutch Cochrane checklist	Bourguet 2007	Sloka 2007	Bourguet 2006	Shie 2008	Selman 2008	Isasi 2005	Magne 2008	Havrilevsky 2005	AHRQ 2006
Adequate search question?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adequate search strategy?	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Adequate study selection?	Yes	Yes	Yes	Yes	Yes	-	Partly	Yes	Yes
Adequate quality appraisal?	Unclear	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Adequate description of data extraction?	No	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes
Description of most important features of included studies?	Yes	Partly	Yes	Yes	Yes	Yes	No	Yes	Yes
Correct meta-analysis?	NA	No	NA	NA	NA	Yes	NA	NA	NA
Valid and applicable systematic review?	**Partly**	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**	**Yes**	**Yes**

Other cancers

Dutch Cochrane checklist	Krug 2008	Van Vliet 2008	Westerterp 2006	Ho Song 2008	Wong 2008	Gillham 2007	Machtens 2007	De Witt 2008	Dong 2008	Czernin 2007
Adequate search question?	Yes	Yes	Yes	Yes	Partly	Yes	Yes	Yes	Yes	Yes
Adequate search strategy?	Yes	Partly	Unclear	Partly	Yes	Partly	NA	Yes	Yes	NA
Adequate study selection?	Yes	Yes	Unclear	Partly	NA	NA	NA	Yes	Yes	No
Adequate quality appraisal?	Yes	Partly	Unclear	No	NA	NA	NA	Yes	Yes	NA
Adequate description of data extraction?	Yes	Yes	No	NA	No	No	No	Yes	Partly	No
Description of most important features of included studies?	Yes	Yes	Yes	Yes	Partly	Yes	Partly	Yes	Yes	Partly
Correct meta-analysis?	Yes	Yes	NA	NA	No	No	NA	NA	Yes	NA
Valid and applicable systematic review?	**Yes**	**Partly**	**Partly**	**No**	**No**	**No**	**No**	**Yes**	**Yes**	**No**

Other pathologies

Dutch Cochrane checklist	SBU 2008	Suchowersky 2006	Beamlands 2007	Nandalur 2008	Termaat 2005
Adequate search question?	Yes	Yes	Partly	Yes	Yes
Adequate search strategy?	Yes	Yes	Yes	Yes	Yes
Adequate study selection?	Yes	Yes	Partly	Yes	Yes
Adequate quality appraisal?	Yes	Yes	Partly	Yes	Yes
Adequate description of data extraction?	Yes	NA	No	Yes	Yes
Description of most important features of included studies?	Yes	Yes	Yes	Yes	NA
Correct meta-analysis?	NA	NA	No	NA	Yes
Valid and applicable systematic review?	**Yes**	**Yes**	**Partly**	**Yes**	**Yes**

DIAGNOSTIC ACCURACY STUDIES

For each tumour or disease, the quality appraisal of all included diagnostic accuracy studies is synthesised in 1 table indicating the number of studies scoring yes, no or unsure on each item of the QUADAS checklist.

Lung cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	17	0	0
Were selection criteria clearly described?	17	0	0
Is the reference standard likely to correctly classify the target condition?	17	0	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	15	0	2
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	17	0	0
Did patients receive the same reference standard regardless of the index test result?	13	4	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	16	0	1
Was the execution of the index test described in sufficient detail to permit replication of the test?	17	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	16	0	1
Were the index test results interpreted without knowledge of the results of the reference standard?	12	0	5
Were the reference standard results interpreted without knowledge of the results of the index test?	10	0	7
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	7	1	9
Were uninterpretable/ intermediate test results reported?	9	2	5
Were withdrawals from the study explained?	12	2	3

Lymphoma

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?			
Were selection criteria clearly described?	2	2	0
Is the reference standard likely to correctly classify the target condition?	3	0	1
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	2	0	2
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	4	0	0
Did patients receive the same reference standard regardless of the index test result?	2	2	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	0	2	2
Was the execution of the index test described in sufficient detail to permit replication of the test?	4	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	0	4	0
Were the index test results interpreted without knowledge of the results of the reference standard?	2	1	1
Were the reference standard results interpreted without knowledge of the results of the index test?	1	2	1
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	0	0	4

Were uninterpretable/ intermediate test results reported?	2	0	2
Were withdrawals from the study explained?	2	1	1

Head and neck cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	8	0	0
Were selection criteria clearly described?	8	0	0
Is the reference standard likely to correctly classify the target condition?	4	3	1
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	3	1	4
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	8	0	0
Did patients receive the same reference standard regardless of the index test result?	4	4	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	6	0	2
Was the execution of the index test described in sufficient detail to permit replication of the test?	8	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	4	3	1
Were the index test results interpreted without knowledge of the results of the reference standard?	7	0	1
Were the reference standard results interpreted without knowledge of the results of the index test?	2	0	6
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	8	0	0
Were uninterpretable/ intermediate test results reported?	5	3	0
Were withdrawals from the study explained?	5	3	0

Colorectal cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	7	1	1
Were selection criteria clearly described?	7	1	1
Is the reference standard likely to correctly classify the target condition?	3	6	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	6	0	3
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	9	0	0
Did patients receive the same reference standard regardless of the index test result?	3	6	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	7	1	1
Was the execution of the index test described in sufficient detail to permit replication of the test?	9	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	7	1	1
Were the index test results interpreted without knowledge of the results of the reference standard?	7	0	2
Were the reference standard results interpreted without knowledge of the results of the index test?	1	1	7
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	9	0	0
Were uninterpretable/ intermediate test results reported?	0	9	0
Were withdrawals from the study explained?	1	8	0

Malignant melanoma

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	4	0	0
Were selection criteria clearly described?	4	0	0
Is the reference standard likely to correctly classify the target condition?	2	2	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	2	0	2
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	4	0	0
Did patients receive the same reference standard regardless of the index test result?	2	2	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	2	2	0
Was the execution of the index test described in sufficient detail to permit replication of the test?	4	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	4	0	0
Were the index test results interpreted without knowledge of the results of the reference standard?	2	0	2
Were the reference standard results interpreted without knowledge of the results of the index test?	0	0	4
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	4	0	0
Were uninterpretable/ intermediate test results reported?	0	4	0
Were withdrawals from the study explained?	1	3	0

Breast cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	7	0	0
Were selection criteria clearly described?	6	0	1
Is the reference standard likely to correctly classify the target condition?	7	0	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	6	0	1
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	7	0	0
Did patients receive the same reference standard regardless of the index test result?	7	0	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	7	0	0
Was the execution of the index test described in sufficient detail to permit replication of the test?	7	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	6	1	0
Were the index test results interpreted without knowledge of the results of the reference standard?	4	0	3
Were the reference standard results interpreted without knowledge of the results of the index test?	4	0	3
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	7	0	0
Were uninterpretable/ intermediate test results reported?	7	0	0
Were withdrawals from the study explained?	7	0	0

Oesophageal cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	6	0	4
Were selection criteria clearly described?	9	1	0
Is the reference standard likely to correctly classify the target condition?	6	4	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	6	0	4
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	10	0	0
Did patients receive the same reference standard regardless of the index test result?	6	4	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	7	1	2
Was the execution of the index test described in sufficient detail to permit replication of the test?	10	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	7	3	0
Were the index test results interpreted without knowledge of the results of the reference standard?	3	1	6
Were the reference standard results interpreted without knowledge of the results of the index test?	1	1	8
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	10	0	0
Were uninterpretable/ intermediate test results reported?	0	10	0
Were withdrawals from the study explained?	1	9	0

Gastric cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	2	0	1
Were selection criteria clearly described?	3	0	0
Is the reference standard likely to correctly classify the target condition?	1	1	1
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	2	0	1
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	3	0	0
Did patients receive the same reference standard regardless of the index test result?	2	1	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	2	0	1
Was the execution of the index test described in sufficient detail to permit replication of the test?	3	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	2	1	0
Were the index test results interpreted without knowledge of the results of the reference standard?	1	0	2
Were the reference standard results interpreted without knowledge of the results of the index test?	1	0	2
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	3	0	0
Were uninterpretable/ intermediate test results reported?	0	3	0
Were withdrawals from the study explained?	0	3	0

Thyroid cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	10	0	0
Were selection criteria clearly described?	10	0	0
Is the reference standard likely to correctly classify the target condition?	2	7	1
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	4	0	6
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	9	0	1
Did patients receive the same reference standard regardless of the index test result?	2	7	1
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	5	1	4
Was the execution of the index test described in sufficient detail to permit replication of the test?	10	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	5	4	1
Were the index test results interpreted without knowledge of the results of the reference standard?	4	0	6
Were the reference standard results interpreted without knowledge of the results of the index test?	0	0	10
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	10	0	0
Were uninterpretable/ intermediate test results reported?	0	10	0
Were withdrawals from the study explained?	0	10	0

Pancreatic cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	4	0	2
Were selection criteria clearly described?	5	1	0
Is the reference standard likely to correctly classify the target condition?	0	5	1
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	0	0	6
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	6	0	0
Did patients receive the same reference standard regardless of the index test result?	0	6	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	1	1	4
Was the execution of the index test described in sufficient detail to permit replication of the test?	6	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	2	4	0
Were the index test results interpreted without knowledge of the results of the reference standard?	6	0	0
Were the reference standard results interpreted without knowledge of the results of the index test?	0	0	6
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	6	0	0
Were uninterpretable/ intermediate test results reported?	0	6	0
Were withdrawals from the study explained?	0	6	0

Primary liver cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	2	0	0
Were selection criteria clearly described?	2	0	0
Is the reference standard likely to correctly classify the target condition?	0	2	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	0	0	2
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	2	0	0
Did patients receive the same reference standard regardless of the index test result?	0	2	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	1	1	0
Was the execution of the index test described in sufficient detail to permit replication of the test?	2	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	2	0	0
Were the index test results interpreted without knowledge of the results of the reference standard?	0	0	2
Were the reference standard results interpreted without knowledge of the results of the index test?	0	0	2
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	2	0	0
Were uninterpretable/ intermediate test results reported?	0	2	0
Were withdrawals from the study explained?	0	2	0

Cervical cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	3	0	0
Were selection criteria clearly described?	3	0	0
Is the reference standard likely to correctly classify the target condition?	0	3	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	0	0	3
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	3	0	0
Did patients receive the same reference standard regardless of the index test result?	0	3	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	0	2	1
Was the execution of the index test described in sufficient detail to permit replication of the test?	3	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	3	0	0
Were the index test results interpreted without knowledge of the results of the reference standard?	3	0	0
Were the reference standard results interpreted without knowledge of the results of the index test?	1	0	2
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	3	0	0
Were uninterpretable/ intermediate test results reported?	0	3	0
Were withdrawals from the study explained?	0	3	0

Ovarian cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	2	0	0
Were selection criteria clearly described?	2	0	0
Is the reference standard likely to correctly classify the target condition?	1	1	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	0	0	2
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	2	0	0
Did patients receive the same reference standard regardless of the index test result?	1	1	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	1	1	0
Was the execution of the index test described in sufficient detail to permit replication of the test?	2	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	2	0	0
Were the index test results interpreted without knowledge of the results of the reference standard?	1	0	1
Were the reference standard results interpreted without knowledge of the results of the index test?	0	0	2
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	2	0	0
Were uninterpretable/ intermediate test results reported?	0	2	0
Were withdrawals from the study explained?	0	2	0

Uterine cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	4	0	0
Were selection criteria clearly described?	4	0	0
Is the reference standard likely to correctly classify the target condition?	2	2	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	1	0	3
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	4	0	0
Did patients receive the same reference standard regardless of the index test result?	2	2	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	2	2	0
Was the execution of the index test described in sufficient detail to permit replication of the test?	4	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	4	0	0
Were the index test results interpreted without knowledge of the results of the reference standard?	2	0	2
Were the reference standard results interpreted without knowledge of the results of the index test?	1	0	3
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	4	0	0
Were uninterpretable/ intermediate test results reported?	0	4	0
Were withdrawals from the study explained?	0	4	0

Renal cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	1	0	0
Were selection criteria clearly described?	1	0	0
Is the reference standard likely to correctly classify the target condition?	1	0	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	1	0	0
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	1	0	0
Did patients receive the same reference standard regardless of the index test result?	1	0	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	1	0	0
Was the execution of the index test described in sufficient detail to permit replication of the test?	1	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	1	0	0
Were the index test results interpreted without knowledge of the results of the reference standard?	1	0	0
Were the reference standard results interpreted without knowledge of the results of the index test?	0	0	1
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	0	0	1
Were uninterpretable/ intermediate test results reported?	0	1	0
Were withdrawals from the study explained?	0	1	0

Testicular cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	2	0	0
Were selection criteria clearly described?	2	0	0
Is the reference standard likely to correctly classify the target condition?	2	0	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	1	0	1
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	1	1	0
Did patients receive the same reference standard regardless of the index test result?	2	0	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	2	0	0
Was the execution of the index test described in sufficient detail to permit replication of the test?	2	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	2	0	0
Were the index test results interpreted without knowledge of the results of the reference standard?	1	0	1
Were the reference standard results interpreted without knowledge of the results of the index test?	1	0	1
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	2	0	0
Were uninterpretable/ intermediate test results reported?	0	2	0
Were withdrawals from the study explained?	0	1	1

Brain cancer

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	5	0	0
Were selection criteria clearly described?	4	0	1
Is the reference standard likely to correctly classify the target condition?	5	0	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	5	0	0
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	5	0	0
Did patients receive the same reference standard regardless of the index test result?	5	0	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	5	0	0
Was the execution of the index test described in sufficient detail to permit replication of the test?	5	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	5	0	0
Were the index test results interpreted without knowledge of the results of the reference standard?	5	0	0
Were the reference standard results interpreted without knowledge of the results of the index test?	1	0	4
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	5	0	0
Were uninterpretable/ intermediate test results reported?	5	0	0
Were withdrawals from the study explained?	5	0	0

Infectious diseases

Item	Yes	No	Unclear
Was the spectrum of patients representative of the patients who will receive the test in practice?	2	0	0
Were selection criteria clearly described?	2	0	0
Is the reference standard likely to correctly classify the target condition?	0	2	0
Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	0	0	2
Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	2	0	0
Did patients receive the same reference standard regardless of the index test result?	0	2	0
Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	0	0	2
Was the execution of the index test described in sufficient detail to permit replication of the test?	2	0	0
Was the execution of the reference standard described in sufficient detail to permit its replication?	0	2	0
Were the index test results interpreted without knowledge of the results of the reference standard?	0	1	1
Were the reference standard results interpreted without knowledge of the results of the index test?	0	0	2
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	2	0	0
Were uninterpretable/ intermediate test results reported?	0	2	0
Were withdrawals from the study explained?	0	2	0

APPENDIX 3: EVIDENCE TABLES

LUNG CANCER

SOLITARY PULMONARY NODULE

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with SPN	**FDG-PET** Mixed reference standard	One SR (DACEHTA 2001), included in previous KCE report.	Good-quality HTA Search date: Aug 2005 Databases: Medline, EMBASE, Cochrane Library, HTA database, DARE, individual contacts through INAHTA No meta-analysis
Systematic reviews				
Wahidi 2007	Patients with SPN (size not mentioned)	**FDG-PET** Comparator: dynamic CT Mixed reference standard	For PET (17 studies, 790 patients from 1990 to 2004): 2 not included in KCE 2005 Orino 1998 (n= 23): Se 88 % (15/17; 95%CI 62%-98%), Sp 67% (4/6; 95%CI 24%-94%) Matthies 2002 (n=36): Se 80% (16/20; 95%CI 55%-93%), Sp 94% (15/16; 95%CI 67%-99%) For dynamic CT (7 studies, 948 patients from 1992 to 2004): Se 98-100%, Sp 54-93%	High-quality SR included in ACCP Evidence-Based Clinical Practice Guidelines (2nd Edition) 2007 Search date: August 2005 Databases : MEDLINE, and the Cochrane Library
Cronin 2008	Patients with SPN (< 30 mm)	**FDG-PET** Comparators: CT, MRI, SPECT Reference standards unclear	PET (22 studies from 1990 to 2005). Three not included in KCE 2005 nor in Wahidi 2007: Hagberg 1997 (n=49): +LR 3 (1-9.1), -LR 0.13 (0.032- 0.44) Lin 2004 (n=15): +LR 1.4 (0.6-3.3), -LR 0.14 (0.02- 0.57) Halley 2005 (n=28): +LR 3.2 (1.1-8.9), -LR 0.08 (0.01- 0.38) All studies combined n= 1.069 : Pooled LR+ : 5.44 (3.56, 7.32) Pooled LR- : 0.06 (0.02, 0.09) SPECT : (7 studies from 1999 to 2005): Combined n= 421 : Positive LR : 5.2 (4-6.3) Negative LR : 0.06 (0.04- 0.08) Positive LRs for others diagnostic tests were: CT 3.91 (95%CI 2.42, 5.40), MRI 4.57 (3.03, 6.1) . Negative	High-quality M-A Search date: December 2005 Databases : MEDLINE, OLDMEDLINE, CANCERLIT,and the Cochrane Library Software: Stata version 9.0

	Population	Index test	Results	Comments
			LRs: CT 0.10 (0.03, 0.16), MRI 0.08 (0.03,0.12) Author's conclusion : no better performance for PET than for SPECT.	

Primary studies

	Population	Index test	Outcome	Results	Comments
Buck 2005	47 patients suspected of pulmonary malignancy on CT	**FLT-PET** Reference standard : histopathology	Performance of FLT-PET for detection of primary lung cancer	FLT-PET: Se: 90%, 95%CI 68%-98% Sp: 100%, 95%CI 75%-100%	Analysis made on 46/47 patients
Hashimoto 2006	43 patients suspected of pulmonary malignancy on CT	**FDG-PET** Reference standard : histopathology or imaging	Performance of 18FDG-PET for diagnose of primary lung cancer (if FDG SUV < 2.5)	FDG-PET if FDG SUV < 2.5 : Se: 81%, 95%CI 54%-95% Sp: 85%, 95%CI 65%-95%	Retrospective study
Ferran 2006	29 patients suspected (with conventional methods) of pulmonary malignancy	**FDG-PET/CT** Comparator: SPECT-CT Reference standard : mixed (surgery, FNA or BAL n = 1 (bronch-alveolair aspiration))	Performance of both for detection of primary lung cancer	PET/CT: Se: 100% (20/20), 95%CI 80%-100% Sp: 89 % (8/9), 95%CI 50%-99% SPECT-CT: Se: 85% (17/20), 95%CI 61%-95% Sp: 89 % (8/9), 95%CI 50%-99%	Prospective study
Wang 2007	44 patients suspected (with conventional methods) of pulmonary malignancy	**FDG-PET** (dual head coincidence imaging) Comparator: SPECT Reference standard : histopathology or clinical follow-up (n=6)	Performance of both for detection of primary lung cancer	18 FDG-PET: Se: 100% (31/31), 95%CI 86%-100% Sp: 46 % (6/13), 95%CI 20%-74% SPECT: Se: 100% (31/31), 95%CI 86%-100% Sp: 69 % (9/13), 95%CI 39%-89%	All patients underwent both imaging (one day after) Discordant data in table 2
Tsunezuka 2007	150 consecutive patients suspected of pulmonary malignancy	**FDG-PET** Reference standard :	Accuracy of 18F-FDG-PET in distinguishing malignancy from	All lesions: Se: 76%, 95%CI 65%-84% Sp: 64%, 95%CI 51%-75%	Study type (prospective or retrospective) not mentioned.

	Population	Index test	Outcome	Results	Comments
		histopathology	benign lesions in small or very small lesions (<2.0-cm diameter).	Lesions < 2 cm: Se: 51%, 95%CI 34%-68% Sp: 52%, 95%CI 32%-71% Author's conclusions: The accuracy of 18F-FDG-PET is generally low in distinguishing malignancy from benign lesions in small lesions (<2.0-cm diameter).	
Kaira 2009	43 patients with lesion suspected of pulmonary malignancy	**FLT-PET** **FDG-PET** Reference standard : histopathology	Evaluation of primary lesions	FLT-PET: Se: 84%, 95%CI 67%-93% Sp: 100%, 95%CI 52%-100% No sufficient data to calculate sensitivity/specificity for FDG-PET, but according to author sensitivity was 89%	Sample size of benign diseases is small (6 patients)

NSCLC

Staging

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	NSCLC patients	**FDG-PET**	2 systematic reviews (Toloza 2003 and HTBS 2001) and 2 primary studies (van Tinteren 2002, Viney 2004). Toloza 2003 and Viney 2004 were not included in the previous KCE report. Mediastinal staging : Toloza 2003: 1 SR : PET Se 84% (95%CI 78-89%), Sp 89% (95%CI 83-93%) 3 primary studies (n=584): PET Se 61–68%, Sp 72–84% Viney 2004: staging to avoid futile thoracotomy (*n* = 183: 91 PET+CWU, 92 to CWU only). Few patients: 4/91 PET+CWU vs 2/92 CWU (*p* = 0.2) avoided thoracotomy. Metastasis : Toloza 2003: PET had sensitivity and specificity of over 90% for detection of any distant metastases apart from the brain (based on 1 primary study, n= 100) CT+MR 100% accurate, PET Se 60%, Sp 99%	See above
NCCHTA	NSCLC patients	**FDG-PET/CT**	3 primary studies:	See above

	Population	Index test	Results	Comments
2007		Reference standard: histology	Antoch (2003), n = 27 PET/CT significantly better than PET or CT. Twenty p/26 correctly staged by PET and 19/26 by CT. This led to changes in treatment plan for four patients (15%). Cerfolio (2004), n = 129 T and M staging: PET/CT better than PET, particularly for staging levels I and II Metastasis: PET/CT changed management in 12 patients (9%), PET changed management in one patient Lardinois (2003), n = 49 T and M staging: PET/CT significantly higher diagnostic accuracy than CT or PET ($p < 0.001$) or PET + CT ($p = 0.01$) Metastasis: PET/CT provided additional information in 20/49 patients (41%) beyond that provided by visual correlation of PET and CT.	

Primary studies

	Population	Index test	Outcome	Results	Comments
Chin A Yi 2007	134 patients with stage T1 NSCLC	**FDG-PET/CT** Comparator: Helical dynamic CT Reference standard : surgery or mediastinoscopy	Mediastinal lymph node staging	Helical dynamic CT Se: 65% (22/34), 95%CI 46%-80% Sp: 89% (97/109), 95%CI 81%-94% FDG PET-CT (with SUV) Se: 56% (19/34), 95%CI 38%-72% Sp: 100% (109/109), 95%CI 96%-100%	Prospective study
Ohno 2007	115 patients with stage I/II NSCLC	**FDG-PET/CT** Comparator: STIR turbo imaging (with lymph node saline ratio LSR) Reference standard : histopathology	N-stage assessment	STIR turbo SE imaging Se: 91% (39/43), 95%CI 77%-97% Sp: 93% (67/72), 95%CI 84%-97% FDG-PET/CT Se: 77% (33/43), 95%CI 61%-87% Sp: 88% (63/72), 95%CI 77%-94%	Prospective study

	Population	Index test	Outcome	Results	Comments
Ohno 2008	203 patients with NSCLC	**FDG-PET/CT** Comparator: MRI with DW or without imaging Reference standard : histopathology or radiology (12 months follow-up)	M-stage assessment	MRI with DW Se: 68% (23/34), 95%CI 49%-82% Sp: 92% (150/163), 95%CI 86%95%] FDG-PET/CT Se: 71% (24/34), 95%CI 52%-84% Sp: 88% (63/92), 95%CI 77%-94%	Prospective study
Quaia 2008	150 patients involved in preoperative diagnosis of mediastinal nodal metastasis in stage T1 NSCLC	**FDG-PET/CT** Comparator: Contrast-enhanced CT Reference standard : histopathology (thoracotomy)	N-staging	Contrast-enhanced CT Se: 46% (23/50), 95%CI 32%-60% Sp: 93% (14/15), 95%CI 66%-99% FDG-PET/CT Se: 78% (39/50), 95%CI 63%-88% Sp: 80% (12/15), 95%CI 51%-94%	Prospective study
Kim 2006	150 patients with stage T1 NSCLC	**FDG-PET/CT** Reference standard: histopathology	Mediastinal node staging	PET/CT Se: 47% (16/34), 95%CI 30-65% Sp: 100% (116/116), 95%CI 97-100%	Prospective study.
Kim 2007	674 NSCLC patients (all stages), referred for surgery	**FDG-PET/CT** Reference standard : histopathology by mediastinoscopy or thoracothomy	Mediastinal node staging	PET/CT Se: 61% (110/180), 95%CI 54-68% Sp: 96% (473/494), 95%CI 94-97%	Prospective study. Possible selection bias.
Lee BE 2007	Patients with biopsy-proven NSCLC (n=336)	**FDG-PET** (n=210) **FDG-PET/CT** (n=126) Reference standard : histopathology by mediastinoscopy or thoracotomy	Mediastinal node staging	PET Se: 61% (43-77%) Sp: 94% (90-97%) PET/CT Se: 86% (67-96%) Sp: 81% (71-88%)	Retrospective study.

Prognosis

	Population	Index test	Results	Comments
Systematic reviews				
Berghmans 2008	Patients with stages I to III/IV NSCLC, prognostic value at diagnosis	**FDG-PET** Outcome: mortality	13 studies : Ahuja1998 (n=155) HR 2.05, 95% CI: [1.24-3.37], Dhital 2000(n=77) HR 1.30, 95% CI: [0.70-2.60], Downey 2004(n=100) HR 2.60, 95% CI: [1.02-6.64], Eschmann 2006(n=137) HR 1.71, 95% CI: [1.00-2.93], Higashi 2002(n=57) HR 6.20, 95% CI: [1.34-28.75], Jeong 2002(n=73) HR 4.33, 95% CI: [1.80-10.45], Port 2005(n=64) HR 2.36, 95% CI: [0.24-22.88], Prevost 2005(n=120) HR 2.36, 95% CI: [1.34-4.15], Sasaki 2005(n=162) HR 7.66, 95% CI: [1.41-41.50], Sugawara 1999(n=38) HR 0.56, 95% CI: [0.21-1.44], Vansteenkiste 1999(n=125) HR 2.72, 95% CI: [1.50-4.94], Borst 2005, (n=51) HR 3.15, 95% CI: [1.59-6.22] Cerfolio 2005(n=315) HR 2.65, 95% CI: [1.63-4.31]. High SUV is identified as a poor prognostic factor for survival. The combined HR of SUV for the 11 reports including patients with confirmed pathologies was 2.07 (95%CI 1.66-2.58) for fixed effects and 2.13 (95%CI 1.54-2.95) for random effects.	Search date : June 2006 Methodological quality of primary studies was low

Primary studies

	Population	Index test	Outcome	Results	Comments
At diagnosis					
Downey 2007	487 patients with R0 resection for NSCLC without induction or adjuvant therapy	**FDG-PET**	Prediction of overall survival	SUV : independent predictor of survival (P = 0.03), adjusting for tumor size (P = 0.02) and histology (P<0 .01). SUV after adjusting for clinical TNM stage : independent predictor of survival (P = 0.03) SUV after adjusting for pathologic TNM stage: not an independent predictor of survival (P = 0.09)	Good quality 33 patients yet included in Berghmans conclusion : added value of FDG-PET remains questionable
Hoang 2008	214 patients advanced-stage NSCLC	**FDG-PET**	Prediction of survival	SUV alone : no statistical difference (p = 0.11) for SUV<11.1 versus SUV >11.1) SUVmax after adjusting for sex, stage and treatment : no statistical difference . If uncategorized values : p = 0.35, if categorized values : p =0.45	Retrospective review on long period with variations in treatment protocols
Tanvetyanon	59 patients with	**FDG-PET**	Prediction of survival	PET response (semiquantitative reading) is not	Data were issued from

	Population	Index test	Outcome	Results	Comments
2008	histologically confirmed NSCLC who had resectable disease, including stages IB, II, IIIA, or IIIB			prognostic of survival (P =0.38). Stage : is significant predictor of survival (hazard ratio = 4.58; 95% CI, 1.24 to 16.73; P =0 .02 Completeness of resection: is significant predictors of survival (hazard ratio = 3.76; 95% CI, 1.29 to 10.92; P = 0 .02).	phase II clinical trial
Shin 2008	184 patients with NSCLC	**FDG-PET/CT** Reference: histopathology	Prediction of overall survival (OS) and disease free survival (DFS)	PET/CT results for N2 disease detection: Se : 48 % (11/23) Sp : 95% (153/161) 3-year DFS rate in the PET/CT FN group: 31%, 95% CI [13.6-48.0%]) 3-year DFS rate in the PET/CT TP group: 16%, 95% CI [1.7-29.5% [(p = 0.649) 3-year DFS rate in the P-TN group: 77%, 95% CI [72.0-81.2%] (p < 0.001)	Good quality
Nguyen 2006	53 patients NSCLC (31stage I, 15 stage II, and 7 stage III disease)	**FDG-PET/CT**	Prediction of DFS	Univariate analysis : (DFS) significantly correlated with maxSUV (<7 versus ≥7, p = 0.001), % Ki-67 expression (<25% versus ≥25%, p = 0.047), tumor size (<3 cmversus ≥3 cm, p = 0.027), and tumor cell differentiation (well/moderate versus poor, p = 0.011). In multivariate analysis including T-classification, age and histology, maxSUV is an independant predictor of recurrence (p=0.002) and death (p = 0.041).	Multivariate Cox proportional analysis (SPSS software) Median follow-up duration : 15 months.
Goodgame 2008	136 patients (T1 = 77, T2 = 59) with resection for NSCLC without induction or adjuvant therapy	**FDG-PET**	Prediction of recurrence and overall survival	Multivariate analysis : including T-classification, age and histology, high SUV is independently associated with recurrence (p= 0.002) and death (p = 0.041).	Retrospective study Consecutive patients but only patients who had preoperative FDG-PET were selected (bias ?). Median follow-up duration: 46 months
After treatment					
Mac Manus 2005	88 patients after concurrent platinum-based radical chemo/RT (n = 73) or radical RT alone (n =	**FDG-PET**	Prediction of OS in four groups : CMR : complete metabolic response	Survival at 2 y after adjusting for pre-treatment status, weight loss and PET stage (multifactorial) : CMR (48) : HR 1.00, Not CMR (40) : HR 2.71, IC 95% [1.58-4.7[(p =	Good study design (prospective) but results based on survival at 2 y and not on all data

	Population	Index test	Outcome	Results	Comments
	15)		PMR : partial metabolic response SMD : stable metabolic disease PMD : progressive metabolic disease	0.0001) Author's conclusion: Attainment of CMR after radical RT/chemoRT for NSCLC bestows superior freedom from local and distant relapse; late local relapse is common.	available (survival at 3 and 4 y)
Hoekstra 2005	47 patients receiving induction chemotherapy (3 cycles in toto). After IC, several different treatments were applied (following EORTC protocol). Median survival : 21 months	**FDG-PET**	Monitoring response is the final objective but results are focused on prognosis value.	Comparison of prognostic value for predicting survival : MLN status (measured by PET) predicted survival (hazard ratio [HR], 2.33; 95% CI, 1.04 to 5.22; P = 0.04) CT: HR, 1.87; 95% CI, 0.81 to 4.30; P =0.14. FDG-PET (MR glu) : HR, 1.95; 95% CI, 1.28 to 2.97 (P =0 .002). Multivariate analysis combining CT and glucose consumption (after 3 cycles) : Residual MRglu after one cycle selected patients with different outcomes (HR, 2.04; 95% CI, 1.18 to 3.52; P =0 .01).	Prospective design Multivariate analysis performed on Cox proportional hazards regression. Results are presented by hazard ratios with 95% CIs. Basic assumptions of linearity and additivity were checked.
Ohtsuka 2006	98 patients NSCLC (63 stage IA, 35 stage IB)	**FDG-PET/CT**	Prediction of DFS	Univariate analysis : SUV (threshold 3.3) : SE : 91.7% , SP : 62.8% (p = 0.008) Histologic grade of differentiation : difference between moderately or poorly differentiated adenocarcinomas and well-differentiated adenocarcinomas was statistically significant (P = 0.036). Multivariate analysis: SUV with a cutoff value of 3.3 did not achieve statistical significance(P =0.079). HR = 4.2 (IC 95% = 0.8-21.5) Histologic grade of cell differentiation was found to have no correlation with tumor recurrence (P = 0.286). HR = 2.4 (IC 95% = 0.5-12.2)	Prospective design Receiver operating characteristic (ROC) curves of SUV for the prediction of recurrence were generated using MedCalc (Medisoftware, Mariakerke, Belgium) by plotting sensitivity versus 1-specificity for varying thresholds of SUV. The best combination between sensitivity and specificity was found.
Recurrent disease					
Hellwig	62 consecutive patients	**FDG-PET**	Prediction of survival	FDG-PET(diagnostic performance) :	Prospective study

	Population	Index test	Outcome	Results	Comments
2006	with suspected recurrence after surgical therapy			SE : 93% (95%CI: 86–100%) SP : 89% (95%CI: 74–100%) Predicting values of SUV : SUV in recurrent tumour : higher than in benign changes (10.6±5.1 vs 2.1±0.6, <0.001). SUV in median survival : if SUV<11: 18 months, if SUV≥11: 9 months, p<0.01 SUV in median survival after surgery : If SUV<11:, 46 months,if SUV≥11: 3 months, p<0.001). SUV in recurrent tumour was identified as an independent prognostic factor (p<0.05).	Multivariate analysis was carried out by Cox regression analysis with backward stepwise exclusion at a significance level of 0.10 to identify prognostic factors in respect of survival. (SPSS program package)

Treatment response

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	NSCLC patients	**FDG-PET**	Six primary studies : Choi (2002), n= 30: no 2x2 table Port (2004), n= 25, treatment response before resection, for N1+N2 disease (compared to N0): 12 N0 disease, PET three FPs 13 N1/2 disease, PET five FNs compared to CT five FPs, six FNs Ryu (2002), n = 26, treatment response before resection (PET): Se: 67%, 95% CI [41%-86%] Sp: 62%, 95% CI [26%-90%] for primary tumors (SUV cut-off defined as 3) Se: 88%, 95% CI [62%-98%] Sp: 57%, 95% CI [20%-88%] for lymph nodes (visual) Se: 58%, 95% CI [29%-83%] Sp: 92%, 95% CI [78%-98%]	See above.

	Population	Index test	Results	Comments
			Cerfolio (2003), n= 34: no 2x2 table Schmücking (2003), PET-CT , n= 34 (response): no 2x2 table Schmücking (2005), PET-CT, n= 32 : RCT comparing (1) CRT then chemo vs (2) chemo then CRT, then surgery PET pre- and post-neoadjuvant therapy (before surgery). CT at the same time. PET CR defined as SUV < 2.5. Primary tumour PET CR in 17; 16 had good path response, 1 FN Lymph nodes PET CR in 10; all had good path response, 5 FPs Survival CR vs no CR *p* = 0.008 for OS 2-year survival 76% vs 20% Weber (2003) : n= 57 : advanced NSCLC patients given palliative chemo. Prediction of response using early PET. 20/28 PET responders had a RECIST response 1/27 PET non-responders had a RECIST response FDG net influx also predicts response	

Primary studies

	Population	Index test	Outcome	Results	Comments
De Leyn 2006	30 patients with stage IIIA-N2 NSCLC mediastinal restaging after induction therapy	**FDG-PET/CT** Comparator: Remediastinoscopy Reference standard: thoracotomy	Operability	PET/CT : Se: 77%, 95%CI 50%-92% Sp: 92%, 95%CI 62%-99% Remediastinoscopy: Se: 29%, 95%CI 41%-68% Sp: 100%, 95%CI 62%-100% Sensitivity significantly better for PET-CT (p<0.0001)	Prospective study All patients underwent thoracotomy
Eschman 2007	70 patients with stage III NSCLC	**FDG-PET** before and after neo-adjuvant radio-chemotherapy (NARCT) Reference standard:	Operability	detection of residual viable primary tumor : Se: 95%, 95%CI 80%-99% Sp: 80%, 95%CI 44%-96% presence of lymph node metastases : Se: 77%, 95%CI 58%-90%	Based on same population than population included in publication of 2006 (see Berghmans above in prognosis)

	Population	Index test	Outcome	Results	Comments
		histopathology (n= 47)		Sp: 68%, 95%CI 46%-84%	
Wong 2007	20 patients treated by radiotherapy	**FDG-PET** Comparator: CT	Local control or failure rate	FDG-PET (SUV) : Se: 94%, 95%CI 71%-100% Sp: 50%, 95%CI 3%-97% CT : Se: 67%, 95%CI 41%-68% Sp: 50%, 95%CI 3%-97%	

Radiotherapy planning

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	NSCLC patients	**FDG-PET**	SR of NICE 2005, including 4 primary studies. Four additional PSs : Bradley (2004), n= 26 ; CT simulation used for RT planning and compared to PET : Two M1 disease found by PET, so 24 went on to have radical RT 14 different GTV (gross tumour volume) (reduced in three, increased in 11) De Ruyscher (2005), n = 44 ; stage I–III NSCLC referred for irradiation of mediastinal nodes : LN selected by PET and localised on CT after visual fusion : 29 patients 61.2 Gy, others 64.8 Gy 11 local recurrence, 18 any recurrence or failure (two nodal failure outside PTV) Schmücking (2003), PET-CT , n= 27 (RT planning) : RT planning PTV 3–21% higher with PET Volume of normal lung receiving > 20 Gy reduced by 5–17% Van Der Wel (2005) : n= 21 pathologically proven N2–3 M0 NSCLC. RT planning using visually fused PET+CT : Nodal GTV CT: 13.7 ± 3.8 cm3 PET: 9.9 ± 4.0 cm3 Mean oesophageal dose CT: 29.8 ± 2.5 Gy PET: 23.7 ± 3.1 Gy 14 plans changed; 11 decreased volume Estimated TCP (tumour control probability.): 12.5% vs 18.3%	See above.

Recurrence

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	NSCLC patients	**FDG-PET/CT** **FDG-PET** Reference standard: biopsy, clinical or imaging follow-up	One PS (Keidar 2004) n = 42 patients with suspected recurrent NSCLC (no evidence of malignancy 6 months before) : PET : Se : 96%, 95% CI (80 to 99%) Sp : 53%, 95% CI (31 to 74%) PET-CT : Se : 96 %, 95% CI (80 to 99%) Sp : 82%, 95% CI (59 to 94%) PET/CT contributed to change in management in 12 (29%) patients : In five identified that FDG uptake was benign and so further investigations were not needed In one precise location of malignant sites was identified, allowing RT In three size/location of radiation field was altered In three additional mets were identified leading to altered radiation field and/or chemo	See above.

SCLC

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients diagnosed with SCLC	**FDG-PET** **FDG-PET/CT** Objective: Staging and restaging Mixed eference standard (histology, conventional imaging or clinical follow-up).	PET (3 studies from 2001 to 2004, 162 patients) Studies included: Bradley 2004, Brink 2004, Kut 2007. Sensitivity (3 studies) = 100%, insufficient data provided to calculate specificity. PET-CT (7 studies) Blum 2004 (staging and restaging), n = 36, Se: 100%, Sp: Not calculated Fischer 2006 (staging and restaging), n = 20, Se: 92%, Sp: Not calculated Kamel 2003 (staging and restaging), n = 42, Se: 93%, 95% CI [66%-100%] Sp: 66%, 95% CI [31%-91%] Fischer 2007, n = 26, treatment response before resection (PET), for primary tumors (visual) Se: 93%, 95% CI [64%-100%]	High-quality HTA Study type : prospective or retrospective

	Population	Index test	Results	Comments
			Sp: 100%, 95% CI [52%-100%] Niho 2007, n = 63 Se: 14%, Sp: Not calculated Pandit 2003, n = 46 Se: 97%, 95% CI [85%-100%] Sp: 78%, 95% CI [56%-92%] Vinjamuri 2008, n = 51 Se: 100%, Sp: Not calculated	
AHRQ 2008	Patients diagnosed with SCLC	**FDG-PET** Outcome: impact on management	3 studies included : Blum 2004: considerable change in management for 17/36 patients Bradley 2004: change in management (RTplanning) for 7/23 patients Kamel 2003 (PET and PET-CT): change in management for 12/42 patients.	See above.
NCCHTA 2007	Patients diagnosed with SCLC	**FDG-PET** Mixed Reference standard	Two studies were included: Brink 2004 and Bradley 2004 (see above).	See above.

MESOTHELIOMA

	Population	Index test	Outcome	Results	Comments
Flores 2006	137 patients with pathologically proven mesothelioma on initial staging evaluation Median follow-up: 24 months	**FDG-PET**	Prediction of survival	Prognostic value for predicting survival : FDG-PET (uptake value) Median survivals were 9 and 21 months for the high (>10) (HR = 1.9) and low (<10) standard uptake value groups, respectively (P = 0.02). Multivariate analysis: high standard uptake value tumors were associated with a 1.9 times greater risk of death than low standard uptake value tumors (P=0.01) Mixed histology carried a 2.9 times greater risk of death than epithelioid histology (P =0.01) Stages III and IV had a 1.8 times greater risk of death than stages I and II (P= 0.05).	Multivariate analysis performed using the Cox proportional hazards method

LYMPHOMA

Diagnosis

No new evidence.

Staging

Study	Population	Index test / Reference standard	Results	Comments
HTA reports				
NCCHTA 2007	HL* and NHL*	**FDG-PET** Ref standard unclear ; Hong : consensus or follow-up Naumann: concordance between PET and CWU; follow-up in discordant cases Sasaki: concordance between PET and CWU; treatment response and follow-up in discordant cases	MSAC (2001) 7 primary studies (n=369) 7 additional primary studies: Delbeke (2002), n=45 (23 NHL), no sens or spec reported Hong (2003), n=30 (26 NHL): nodal lesions: sens : PET ≥93% ; CT ≥93% ; Ga 26% extranodal lesions : sens PET 88%, CT 88%, Ga 38% spec 100% for all modalities Jerusalem (2001), n=42 with low-grade NHL: no sens or spec reported Naumann (2004), n=88 with early HL, no sens or spec reported Sasaki (2002), n=46 (42 NHL), nodal lesions: sens CWU*+ PET: 100%; CWU+Ga: 74% extranodal lesions : sens CWU+PET : 95% ; CWU+Ga : 74% Shen (2002), n=30 HL or NHL, sens PET 96%; Ga 72% (outcome not specified) Yamamoto (2004), n=28 NHL, no sens or spec reported, analysis not per patient	Confidence intervals not given
Systematic reviews				
Kwee 2008	HL and NHL	**FDG-PET** **FDG-PET/CT**	PET: Stumpe (1998), La Fougere 2006. Only lesion-based results PET/CT: La Fougere (2006). Only region-based results	
Pakos 2005	HL and NHL	**FDG-PET** Evaluating bone marrow infiltration	13 studies, n=587 Naumann (2004), n=88 HD Elstrom (2003), n=105 HD NHL Hoffmann (2003), n=21 NHL Hong (2003), n=30 HD, NHL Sasaki (2002), n=30 HD, NHL Montravers (2002), n=7 HD Wirth (2002), n=39 HD, NHL Jerusalem (2001), n=42 NHL Jerusalem (2001), n=33 HD Buchmann (2001), n=52 HD, NHL	SROC curve not further specified

Study	Population	Index test / Reference standard	Results	Comments
			Partridge(2000), n=24 HD Moog (1998), n=78 HD, NHL Carr (1998), n=38 NHL SROC curve, sensitivity 51% (95% CI 38-64), specificity 91% (95% CI 85-95)	

Primary studies

	Population	Index test	Outcome	Results	Comments
Tsujikawa 2008	42 patients with histologic diagnosis of NHL who were never treated earlier or relapsed and never received treatment within 6 months	**FDG-PET** Ref standard: WHO classification	Distinction aggressive vs. indolent type	FDG-PET not corrected SUV, positive ≥9,5: Sens 80,7% (60,7-93,5) Spec 81,3% (54,4-96) Partial volume corrected SUV, positive ≥11,2 : Sens 80,7% (60,7-93,5) Spec 62,5% (35,4-84,8)	
Bucerius 2006	42 patients with histologically proven HD or NHL	**FDG-PET** Ref standard: concordant results with conventional imaging=true positive or negative; discordant results further investigated with histology and/or follow-up	Staging	Sens 100% (90,3-100) Spec 100% (54,1-100)	upstaged 15,5%; downstaged 27,5%
Pelosi 2008	194 patients with HL or aggressive NHL	**FDG-PET** Ref standard: Bone marrow biopsy; in case PET+/BMB- further investigation with MRI, targeted biopsy or second PET at end of treatment	Diagnosis of bone marrow disease	Sens 65,3% (50,4-78,3) Spec 98,6% (95,1-99,8)	

Restaging / monitoring treatment response

Study	Population	Index test / Reference standard	Results	Comments
HTA reports				
NCCHTA 2007	HL and NHL after one or more cycles of chemotherapy or CRT	**FDG-PET**	9 primary studies for midtherapy response assessment : Becherer (2002), n=16 (10 HL), after ICT*: excluded <20 patients Cremerius (2002), n=24 NHL, after 3 ICT cycles: results not given Filmont (2002), n=43 (12 HL): results not given Haioun (2005), n=90 aggressive NHL: 2 year survival in patients with PET – after 2 cycles ICT: 90%, PET+: 61% Hutchings (2005), n=85 HL, progression free survival in patients with PET-: 95.2%, in patients with PET+ 59.1% Schot (2003), n=46 recurrent lymphomas (33 NHL), progression over 2 years: persistent uptake after ICT 2.6 (1.0-6.9) Spaepen (2002), n=70 aggressive NHL, after 3-4 cycles, remission after follow-up of 1107 days: PET- 83.8% (95% CI 68.0-93.8), PET+: 0% (0-10.6) Torizuka (2004), n=20, prediction of clinical response at end of therapy: PET < 60%: 90.9% (58.7-99.8); PET> 60%: 0% (0-33.6) prediction of progression free survival: PET-: 50.0% (95% CI 6.8-93.2), PET+: 12.5% (95% CI 1.5-38.3) Zijlstra (2003), n=26: prognostic value of PET after 2 cycles better than that of Gallium scan. PET-: 25% progression-free survival, PET+: 64% progression-free survival	
Systematic reviews				
Kwee 2008	HL and NHL	**FDG-PET** **FDG-PET/CT**	PET : Meany (2007); Bjurberg (2006); Zinzani (2006); Rigacci (2005); Filmont (2004) ; Dittmann (2001) ; Filmont (2003) ; Mikhaeel (2000) ; La Fougere (2006) ; Hernandez-Pampaloni (2006) ; Reinhardt (2005); Freudenberg (2004); Mikosch (2003); Mikhaeel (2000); Bangerter (1999); Bangerter (1999); Stumpe (1998) For Hodgkin disease: Sens 86.2-100%; spec 57.1-100.0% For Non-Hodgkin disease: Sens 60-87.0%, spec 80-100%. For mixed populations: Sens 71.4-100%; spec 86.2-94.5% PET/CT fusion: Schaefer (2007), La Fougere (2006), Rhodes (2006), Freudenberg (2004). One study included HD patients only, the other 3 studies included HD/NHL patients. Sens 92.9-100%, spec 90.6-100.0%	

Primary studies

	Population	Index test	Outcome	Results	Comments
Bucerius 2006	103 histologically proven HD or NHL	**FDG-PET** Ref standard: concordant results with conventional imaging=true positive or negative; discordant results further investigated with histology and/or follow-up	Staging	Sens 68,6% (50,7-83,2) Spec 89,7% (79,9-95,8)	

Recurrence

Primary studies

	Population	Index test	Outcome	Results	Comments
Bucerius 2006	48 patients with histologically proven HD or NHL	**FDG-PET**	Detection of recurrence	Sens 97,5% (86,8-99,9) Spec 75% (34,9-96,8)	Concordant results with conventional imaging=true positive or negative; discordant results further investigated with histology and/or follow-up

Post- treatment evaluation

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	HL and NHL after one or more cycles of chemotherapy or CRT	**FDG-PET**	HTBS HTA report : 8 studies to assess active residual disease sens PET 75-80%, CT 75-80% ; spec PET 90%, CT 45%	See above.
Systematic reviews				
Zijlstra 2006	HL and NHL	**FDG-PET**	Residual disease: For Hodgkin disease: Pooled sensitivity 84% (95% CI 71-91.92%), pooled specificity 90% (95% CI 84-93.94%). For Non Hodgkin disease: pooled sensitivity72% (95% CI 61-82%), pooled specificity 100% (95% CI 97-100%).	
Terasawa 2008	HL and NHL	**FDG-PET**	For Hodgkin disease: Posttherapy evaluation (irrespective of restaging results): sensitivity 50-100%, specificity 67-100%. Residual mass evaluation: sensitivity 43-100%, specificity 67-100%. For Non Hodgkin lymphoma: Posttherapy evaluation (irrespective of restaging results): sensitivity 33-77%, specificity 82-100% respectively. Residual mass evaluation: sensitivity 33-87%, specificity 75-100%.	

Prognosis

Study	Population	Index test / reference standard	Results	Comments
HTA reports				
NCCHTA 2007	HL and NHL	**FDG-PET** Mid treatment	Prediction of relapse: De Wit (2001), n=33: PET: sens 100% (69.1-100.0); spec: 78.3% (56.3-92.5) CT: sens 70.0% (34.8-93.3), spec: 26.1% (10.2-48.4) ESR: sens 50.0% (18.7-81.3), spec: 69.8% (47.1-86.8) Zinzani (2002), n= 56 with bulky (>5 cm) abdominal disease (13 HL). PET: sens 81.8% (48.2-97.7), spec 91.1% (78.8-97.5) CT: sens 81.8% (48.2-97.7), spec 8.9% (10.2-48.4)	
NCCHTA 2007	HL and NHL	**FDG-PET** After treatment	Prediction of relapse: Jerusalem (2003), n=36 HL in patients with CT + for residual mass : PET sens 100% (95% CI 16-100), spec 82% (95% CI 57-96) in patients with CT – for residual mass: PET sens 100% (95% CI 29-100), spec 79% (95% CI 49-95) Naumann (2001), n=42 HL + 15 NHL with CT + for residual mass sens 100% (95% CI 2.5-100), spec 93% (95% CI 80-98) Panizo (2004), n=29 HL with CT + for residual mass ≥2 cm sens 100% (95% CI 66-100), spec 85% (95% CI 62-97) Prediction of progressive disease: Friedberg (2004), n=36 HL, PET+: sens 80% (95% CI 28-99); Ga+ sens 40% (95% CI 5-85) ; spec not given Juweid (2005), n=54 NHL given chemotherapy, only correlation with progression free survival reported Lavely (2003), n=20 HL + 20 NHL after chemotherapy or CRT, results for subgroups according to treatment only	

Primary studies

	Population	Index test	Outcome	Results	Comments
Janikova 2008	99 patient with follicular lymphoma	**FDG-PET** (positive: SUV ≥2,5) Ref standard: Analysis of all scans, histologic and surgical data from nodal and extranodal lesions, and in conjunction with treatment response, patient outcomes and repeated imaging during follow-up	Relapse	Post-treatment prediction of relapse Sens 42,4% (25,5-60,8); spec 90% (79,5-96,2)	
Gallamini 2007	260 newly diagnosed, advanced-stage HL patients	**FDG-PET** after 2 cycles of chemotherapy	Progression-free survival	Prediction of progression free survival multivariate regression including PET, extranodal disease, bulky disease and IPS PET after 2 cycles: p 0.0001, HR 43.0 (20.2 to 91.3); Age >45 years: p 0.046, HR 0.49(0.25 to 0.99); Stage IV disease: p 0.001, HR 2.52 (1.35 to 4.68)	
Schot 2007	101 patients with histologically proven relapse or progression of either aggressive NHL or HL who were intended to be treated with second-line chemotherapy followed by myeloablative therapy and ASCT	**FGD-PET**	Failure-free survival	Prediction of failure free survival multivariate regression including histology, LDH, clinical risk score, and FDG-PET response Clinical risk score HR 1.95 (1.36-2.79) p 0 .001; PET CR HR 0.16 (0.07-0.37) p 0.001; PET PR HR 0.38 (0.21-0.67) p 0.001 (PET NR as reference category)	

HEAD AND NECK

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with suspicion of head and neck cancer: squamous cell carcinoma of upper aerodigestive tract, including oral cavity, nasopharynx, oropharynx, hypopharynx and larynx	**FDG-PET** Comparators: CT, MRI Reference standard: NR (Vermeersch 2003), neck dissection or biopsy of suspicious areas (Khan 2004)	**One systematic review** (Vermeersch 2003): included in previous KCE report. **One primary study,** Khan 2004 (n=44 patients): Head and neck cancer patients with uncertain clinical evaluation (involving other imaging) PET vs neck dissection or biopsy of suspicious areas Se: 92% (95% CI: 72% - 99%) Sp: 65% (95% CI: 41% - 84%) *Conclusion*: Morphological imaging such as CT/MRI is irreplaceable to determine the extension of the tumour in adjacent structure, but may lack specificity 'Where doubt exists' PET may be used to improve specificity of CT/MRI *Recommendation*: PET cannot currently replace CT/MRI because of the need for anatomical localisation, but may be helpful to improve specificity of CT/MRI.	High quality HTA Search date : August 2005 Databases: MEDLINE, MEDLINE in-process and other non- indexed Citations, EMBASE, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effectiveness (DARE) and the HTA database (International database of HTA reports). Individual contacts with members of INAHTA
	Patients with cervical lymph-node metastases	PET for detection of synchronous primaries Reference standard: histopathology, clinical or radiographic follow-up	The same **systematic review** and one additional primary study (Nishiyama 2005) showed that PET could detect some, but not all synchronous primaries that other methods failed to detect. **Nishiyama 2005** (n=53 patients): Newly diagnosed head and neck cancer given clinical exam (endoscopy, CT, CXR, neck/abdomen US) - Detection of synchronous primary tumour Reference standard: pathology after biopsy or 6-month follow-up No data about TP, FP, TN, FN, Se and Sp were provided.	
		PET for diagnosis of occult primary tumour	Two systematic reviews (BCBS 2000; MSAC 2001): both included in previous KCE report. Two additional studies (Miller 2005; Stoeckli 2003) showed that	Small studies with a variety of comparators

	Population	Index test	Results	Comments
		Reference standard: histopathology	PET can detect occult primary tumours in patients with cervical lymph-node metastases. Miller (n=26): Se 67% (35-90%), Sp 93% (66-100%) Stoeckli (n=18): Se 63% (24-91%), Sp 90% (55-100%)	No pooled analysis of sensitivity
NCCHTA 2007	Detection of occult primary tumour of the head and neck in patients with cervical LN metastases	**FDG-PET/CT** Comparators: PET, CT Reference standards: histopathology (n=14), clinical follow-up (n=7)	One prospective study (Freudenberg 2005, n=21) showed that PET/CT detected one more occult primary tumour (12 out of 21) than PET alone or PET+CT. PET (Se 52%) and PET/CT (Se 57%) were both more sensitive than CT (Se 23%). *Recommendation*: PET/CT is a valuable tool to detect occult primary tumour of the head and neck in patients with cervical LN metastases.	Only positive patients with cervical metastases are considered for detecting occult primary tumour

Detection of unknown primary tumour

	Population	Index test	Results	Comments
Systematic reviews				
Dong 2008	Patients with unknown primary tumour	**FDG-PET** **FDG-PET/CT**	Twenty-eight primary studies identified (involving a total of 910 patients). FDG-PET: (n=21) Sensitivity: 78% (95%CI 72-84%) Specificity: 79% (95%CI 74-83%) FDG-PET detected 29% of the tumours that were not detected with conventional imaging. FDG-PET/CT: (n=8) Sensitivity: 81% (95%CI 74-87%) Specificity: 83% (95%CI 78-87%) FDG-PET/CT detected 31% of the tumours that were not detected with conventional imaging. The sensitivities and specificities of the individual studies are provided in the article.	High-quality systematic review *Search date*: September 2007 *Databases*: Medline, EMBASE, Cancerlit Meta-analysis performed

Primary studies

	Population	Index test	Outcome	Results	Comments
Johansen 2008	Patients with neck node metastases from a suspected primary arising from the head and neck region (n=64)	**FDG-PET** (n=42) **FDG-PET/CT** (n=22) Standard: Histopathology or follow-up	Detection of primary tumour	*Sensitivity* 86% (18/21; 95%CI 64-97%) *Specificity* 69% (27/39; 95%CI 52-83%) Change in management in 15 patients (25%).	Prospective study. Differential verification. Four patients excluded from analysis.
Garin 2007	Patients with a carcinoma of unknown primary confirmed histologically for which the primary tumour could not be identified by clinical or biological assessment or by conventional imaging (n=51)	**FDG-PET** Standard: Histology or clinical, biological, and radiological follow-up	Detection of primary tumour	*Sensitivity* 100% (12/12; 95%CI 74-100%) *Specificity* 97% (38/39; 95%CI 87-100%) Change in management in 12 patients (24%).	Prospective study. Differential verification. Unclear which imaging test were included during follow-up.
Freudenberg 2005	Patients with histologically or cytologically proven cervical lymph node metastases of unknown primary head and neck tumour (n=21)	**FDG-PET** **FDG-PET/CT** (contrast-enhanced) Comparator: CT Standard: Histopathology or follow-up	Detection of primary tumour	*Sensitivity*: PET 79% (11/14; 95%CI 49-95%) PET/CT 86% (12/14; 95%CI 57-98%) CT 36% (5/14; 95%CI 13-65%) *Specificity*: PET 71% (5/7; 95%CI 29-96%) PET/CT 100% (7/7; 95%CI 59-100%) CT 57% (4/7 ; 95%CI 18-90%) (PS: data discordant with those provided in article)	Retrospective study. Differential verification.

Staging

	Population	Index test	Results	Comments
HTA reports				
MSAC 2008	Patients with biopsy proven, clinical stage T3/4 Nx or Tx N+ newly diagnosed or recurrent carcinoma of the head and neck	**FDG-PET** **FDG-PET/CT** Reference standard: Pathologic confirmation (histopathology or cytopathology), or clinical follow-up of at least six months Comparator: panendoscopy with or without biopsy, CT and optional MRI CWU: chest radiography, abdominal US, whole-body skeletal Scintigraphy, MRI of head and neck for locoregional staging, CT at sites in question if CWU or	Include 6 HTA reports published between 2001 and January 2008 (MSAC 2001, AETMIS 2001, ICES 2004, KCE 2005, AHTAPol 2006, UK-NCCHTA 2007) + 3 primary studies that investigated the additional value of PET in staging of primary head and neck tumours (Liu 2007; Murakami 2007; Ng 2006). Two fair quality studies reported the accuracy of the addition of PET to staging of lymph node metastases on a per-patient basis (Ng 2006, Murakami 2007). In **Ng 2006** (n=134 patients with SCC in oral cavity), PET was used in addition to CT/MRI. - CT/MRI: Se 31.4% (16.9%–49.3%); Sp 91.9% (84.7%–96.4%) - CT/MRI + PET: Se 57.1% (39.4%–73.7%); Sp 96.0% (90.0%–98.9%) Sensitivity increased with no significant change in specificity. In **Murakami 2007** (n=23 HNC patients prior to tumour resection), PET/CT was used in addition to conventional imaging (endoscopy, CT, MRI). Sensitivity remained unchanged at 90% while specificity increased from 75% to 94% (no precision about TP, FP, FN and TN). In **Liu 2007** (n=300 patients with nasopharyngeal cancer), PET was used in addition to conventional work-up (CWU). - CWU: Se 33% (95% CI: 21–46%); Sp 97% (95% CI: 94–99%) - CWU + PET: Se 84% (95% CI: 72–92%) ; Sp 94% (95% CI: 90–96%) The Australian study (Scott 2007) indicated that PET led to a change in management plans in 32% (95% CI: 20–46%) of 56 patients. PET detected additional lesions in 36% of patients (20/56).	Liu et al. (2007): Fair quality, limited applicability

	Population	Index test	Results	Comments
		FDG–PET suggestive for distant metastases	Of those in whom additional lesions were detected, treatment plans changed in 70% (14/20; 95% CI: 46–88%), and of those with no additional lesions detected, treatment plans changed in 11% (4/36; 95% CI: 3–26%) of patients ($p < 0.001$). Conclusion: Compared to conventional work-up, PET improved pre-treatment staging. PET may more accurately define the locoregional extent of disease and better detect distant metastases that would render the disease incurable.	
NCCHTA 2007	Regional lymph-node involvement in patients with cytologically or histologically proven primary head and neck cancer	**FDG-PET** Comparators: CT, MRI Reference standard: histopathology from neck dissection	Three systematic reviews (Goerres 2003, Vermeersch 2003, BCBS 2000): all included in previous KCE report. 12 additional studies considered PET in staging regional lymph-node involvement: - Clinically N0: 4 studies showed that PET sensitivity was much lower than that of SLNB. - 8 studies in populations of mixed (T1-T3) or unspecified stage patients showed that PET or PET + CT had sensitivity of approximately 80% and specificity of 80–97%. This was comparable to or better than CT or MRI in most studies. One of these studies (Kovacs 2004) used SLNB on PET negative necks to improve sensitivity. This combination reduced the number of radical neck dissections from 45 out of 62 compared with 35 out of 62 on CT. *Conclusion*: PET is a valuable tool for staging and could be complemented with SLNB if PET negative in order to avoid futile radical neck dissections.	See above sROC analysis not performed
NCCHTA 2007	Patients with head and neck cancer	**FDG-PET/CT** **FDG-PET** Reference standard: biopsy, clinical and imaging follow-up Comparators: PET, CT	Four primary studies of PET/CT in various stages of head and neck cancer (Branstetter 2005, Rödel 2004, Schöder 2004, Zanation 2005) showed that PET/CT had slightly higher accuracy than PET, by about 10% (sometimes higher sensitivity, sometimes higher specificity). All studies only presented a per-lesion-analysis.	
Systematic reviews				

	Population	Index test	Results	Comments
Kyzas 2008	Patients with head and neck squamous cell carcinoma (HNSCC)	**FDG-PET** Comparators: conventional diagnostic methods (CT, MRI, and US/FNA) Reference standard: pathology	32 studies (1 236 patients) were considered for meta-analysis. Diagnostic accuracy of FDG PET *32 studies with data on lymph node metastases* Se: 79% (95% CI = 72% to 85%) Sp: 86% (95% CI = 83% to 89%). *For cN0 patients only* Se: 50% (95% CI = 37% to 63%) Sp: 87% (95% CI = 76% to 93%). Comparisons with CT, MRI, CT/MRI, and US/FNA (24 studies) *For all patients* Se: FDG PET (80%; 95% CI = 72% to 87%), others (75%; 95% CI = 65% to 83%) Sp: FDG PET (86%; 95% CI = 82% to 90%), others (79%; 95% CI = 72% to 85%) *For cN0 patients only* Se: FDG PET (52%; 95% CI = 39% to 65%), others (45%; 95% CI = 25% to 67%) Sp: FDG-PET (93%; 95% CI = 87% to 96%), others (87%; 95% CI = 72% to 95%) Conclusion: 18 F-FDG PET has good diagnostic performance in the overall pre-treatment evaluation of patients with HNSCC but still does not detect disease in half of the patients with metastasis and cN0. So, there is little evidence to support the routine use of FDG PET to evaluate possible lymph node metastasis among patients with HNSCC and a clinically negative neck.	MEDLINE search (last update July 31, 2007). No language restrictions Hierarchical regression model WinBUGS software version 1.4 and Intercooled Stata version 8.2 (Stata Corp, College Station, TX).

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Krabbe 2008	38 patients with a newly diagnosed SCC of the oral cavity or oropharynx without signs of cervical lymph node metastasis in the physical examination (clinical N0-neck)	**FDG-PET** Comparators: CT (n=19), MRI (n=10), ultrasonography-guided fine needle aspiration cytology (USgFNAC) (n=5) or US (n=4) Reference: histology with neck dissections or follow-up	Diagnostic properties of FDG-PET to detect the presence of micrometastases in the lymph nodes of clinical N0 necks	In 30 patients a neck dissection was performed after the PET study. The 8 patients without neck dissection had a median follow-up of 3.8 years. There was evidence of a positive neck on histology in 7 patients and by follow-up in 1 patient (8/38, 21%). FDG-PET: Se 50% (95%CI: 21%–78%) Sp 97% (95%CI: 83%–99%) Conventional imaging: Se 50% (95%CI: 21%–78%) Sp 70% (95%CI: 52%–83%) Although FDG PET performed better than conventional imaging modalities, sensitivity was lower than desired. As a consequence, clinical application of FDG PET in the patient staged as N0 is limited.	Prospective study
Ng 2008	160 patients with SCC of the oropharynx or hypopharynx	**FDG-PET** Comparator: multi-detector row computed tomography (MDCT) Reference standard: histopathology or clinical exam or imaging follow-up	Diagnostic properties of FDG-PET to detect the presence of distant malignancies	26 of 160 patients were found to have distant malignancy (16%). FDG-PET: Se: 77% (95%CI: 56%–91%) Sp: 94% (95%CI: 88%–97%) MDCT: Se: 50% (95%CI: 30%–70%) Sp: 98% (95%CI: 93%–99%) PET+MDCT: Se: 81% (95%CI: 61%–93%) Sp: 98% (95%CI: 95%–99%) FDG PET appeared to have acceptable	Large prospective study Possibility of verification and review bias: unclear

	Population	Index test	Outcome	Results	Comments
				diagnostic yield for detection of distant malignancies in patients with newly diagnosed oropharyngeal and hypopharyngeal SCC, but sensitivity remains lower than desired.	
Senft 2008	92 patients with HNSCC, who are candidates for curative treatment and at increased risk for distant metastases (i.e., ≥ 3 lymph node metastases, bilateral lymph node metastases, lymph node metastases of ≥ 6 cm, low jugular lymph node metastases, regional tumour recurrence and second primary tumours assessed by palpation, CT, MRI, and/or USFNA	**FDG-PET** Comparator: Chest CT Reference standard: clinical diagnostic work-up between presentation until a follow-up of 12 months including histopathology	Diagnostic properties of FDG-PET to detect the presence of distant malignancies	In 21% (19/92) of the patients, distant metastases were detected during screening at initial presentation Detection of distant metastases PET: Se 53% (95%CI: 39%–67%) Sp 93% (95%CI: 86%–97%) CT: Se 37% (95%CI: 24%–52%) Sp 95% (95%CI: 88%–98%) PET + CT: Se 63% (48%–76%) Sp 95% (88%–98%) Detection of distant metastases and synchronous second primary tumours PET: Se 58% (95%CI: 45%–70%) Sp 93% (95%CI: 86%–97%) CT: Se 39% (95%CI: 28%–53%) Sp 94% (95%CI: 87%–98%) PET + CT: Se 66% (52%–77%) Sp 94% (87%–98%) Conclusion: FDG-PET is a valuable diagnostic tool in screening for distant metastases in HNSCC patients with high risk factors. Screening with a combination of CT-scan of the thorax and whole body FDG-PET decreases overtreatment. It results in a reduction of futile mostly extensive treatments in these patients.	Prospective study

Restaging

	Population	Index test	Results	Comments
HTA reports				
MSAC 2008	Patients in whom post-therapy evaluation has identified a structural abnormality suspicious for residual head and neck carcinoma	**FDG-PET** in addition to clinical examination and/or CT PET is used as a triage test to reduce the number of additional invasive diagnostic procedures. <u>Reference standard</u>: Biopsy or follow-up or neck dissection	Include 6 HTA reports published between 2001 and January 2008 (MSAC 2001, AETMIS 2001, ICES 2004, KCE 2005, AHTAPol 2006, UK- NCCHTA 2007) + 3 primary studies which reported the utility of PET for the assessment of suspected residual carcinoma and response to therapy. <u>Triage to biopsy (PET scan negative/CT positive)</u> Five fair quality studies (Andrade 2006a, 2006b; Yao 2005, 2007a, 2007b) indicated that PET has a good sensitivity (83% - 100%) and a moderate specificity (68% - 100%) in patients with suspected residual HNC after definitive treatment. In patients who have suspected residual disease following definitive treatment, a negative PET result is highly likely to indicate the absence of disease. **Andrade 2006:** Detection of residual disease with PET/CT *Clinical Exam and CT show suspected residual disease (n=11)* Se 83% (36% - 99%) Sp 100% (46% - 100%) *Clinical Exam or CT show suspected residual disease (n=23)* Se 83% (51% - 97%) Sp 100% (68% - 100%) **Yao 2007**: detection of residual disease vs histopathology: Se 100% (46.3% - 100%) Sp 68.4% (43.5% - 86.4%) SUV <3.0 (PET negative): Se 100% (46.3% - 100%) Sp 84.2% (59.5% - 95.8%) A single Australian study (Ware 2004) reported the impact of PET on treatment intent in 53 patients with suspected residual	Much of the data are retrospective and two of the studies are from the same centre. Andrade assessed performance of PET/CT for treatment response at 8 weeks after the radiation therapy to the HNSCC Yao 2005: Unclear in the text when talking about hemi-necks and/or patients Yao 2007: Analysis based on number of hemi-necks rather than patients (one patient had bilateral disease)

	Population	Index test	Results	Comments
		PET may also identify distant metastases, changing treatment intent from curative to palliative	disease. PET findings changed management in 21 patients (40%). When the PET scan was negative, the most common change was avoidance of surgery, as reported in 88% of patients (15/17). Detection of distant metastases UK-NCCHTA (2007): See below	
NCCHTA 2007	Patients with head and neck cancer	**FDG-PET** Comparators: CT, MRI Reference standard: Histopathology OR clinical follow-up, sometimes with histopathology of lesions obtained by biopsy or surgery	Two systematic reviews (Vermeersch 2003, Goerres 2003): both included in previous KCE report. 7 additional primary studies (Conessa 2004, Kubota 2004, Kunkel 2002, Yao 2004, Goerres 2004, Porceddu 2005, Ware 2004) showed PET Se ≥ 80% and Sp ≥ 90%, which was higher than CT/MRI for restaging. Another systematic review (MSAC 2001): also included in previous KCE report.	See above No distinction between restaging and recurrence

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Horiuchi 2008	31 patients with HNSCC received a concurrent chemotherapy (CCR)	**FDG-PET** Comparators: CT and MRI Reference standard: Histopathology	Post-treatment evaluation (delay: 1 month) Diagnostic performance of PET for restaging after chemo- and radiotherapy	The results of pathological examinations after CCR showed 6 residual cases and 25 ones with a pathologically complete response (pCR) CT + MRI: Se: 83% (95% CI : 36%-99%) Sp: 48% (95% CI : 28%-68%) FDG-PET: Se: 67% (95% CI : 24%-94%) Sp: 80% (95% CI : 59%-92%) Conclusion : FDG-PET has a high specificity but limited sensitivity to discriminate residual cancer from fibrosis or scar at 4 weeks after CCR. FDG-PET at 4 weeks after CCR was too early to perform because of limited sensitivity The timing of post-treatment FDG-PET plays an important role to evaluate the response accurately	Horiuchi assessed performance of PET for treatment response at 4 weeks after the therapy

Detection of recurrence

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with suspicion of recurrent head and neck cancer	**FDG-PET** Comparators: CT, MRI Reference standard: Histopathology or clinical follow-up, sometimes with histopathology of lesions obtained by biopsy or surgery	Two systematic reviews (Vermeersch 2003, Goerres 2003): both included in previous KCE report. 7 additional primary studies (Conessa 2004, Kubota 2004, Kunkel 2002, Yao 2004, Goerres 2004, Porceddu 2005, Ware 2004) showed that PET sensitivity was approximately 80%, with specificity at least 90%, which was more accurate than CT/MRI for recurrence. Another systematic review (MSAC 2001): also included in previous KCE report.	See above No distinction between restaging and recurrence
Systematic reviews				
Brouwer 2008	Patients with suspicion of recurrent laryngeal carcinoma after radiotherapy	**FDG-PET** Comparators: CT, MRI, thallium 201 scintigraphy Reference standards: biopsy taken during direct laryngoscopy and clinical follow-up of 12 months	8 articles on 18FDG-PET were included; 3 of them comprised a comparison with CT and/or MRI (n = 181); 1 study was a case-control study (Stokkel 1998) no eligible studies on CT, MRI, and 201Tl scintigraphy (other than laryngeal carcinoma) FDG-PET (7 studies): Se (95% CI): 89% (80%-94%) Sp (95% CI): 74% (64%-83%) FDG-PET vs CT (Greven 1997; n = 23): Se: 80% vs 58% Sp: 81% vs 100% *Recommendation*: Since the chance of missing a recurrence outweighs the risk of a futile direct laryngoscopy, a sensitive strategy will mostly be used in daily clinical practice.	High quality SR Search period: January 1990 until April 2006 Databases : Medline and Embase Languages: English, German, French, Dutch Articles retrieved were published between 1995 and 2004 Meta-analysis with random-effects models Analysis with MetaDisc software (version Beta 1.1.0)
Isles 2008	Patients with head and neck cancer following radiotherapy or chemoradiotherapy	**FDG-PET** Reference Standards: histology from biopsy	Systematic review including 27 primary studies Accuracy of PET for primary site recurrence/residual disease: Se: 94% (95% CI, 87–97%)	Cochrane, MEDLINE and PubMed electronic databases Date search: 31 October 2007

	Population	Index test	Results	Comments
		or surgical specimen and length of disease free survival	Sp: 82% (95% CI, 76–86%) Accuracy of PET for recurrence / residual disease of nodal metastasis Se: 74% (95% CI, 50–89%) Sp : 88% (95% CI, 74–95%) Conclusion: PET is highly accurate for detecting recurrent or residual head and neck squamous cell carcinoma following chemoradiotherapy. A negative PET scan is highly predictive of the absence of disease, so PET may have the potential to obviate the requirement for planned neck dissections or surveillance endoscopies. Sensitivity of PET scanning is decreased if the interval between treatment and scan is less than 10 weeks. Authors propose an algorithm to orientate physicians in their diagnostic strategy.	Meta-analysis with random-effects models
Liu 2007	Patients with nasopharyngeal carcinoma	**FDG-PET** Comparators: CT, MRI Reference standards: histopathologic analysis and/or close clinical and imaging follow-up for at least 6 months	Systematic review including 21 articles (33 studies; n = 1 813 patients) Diagnostic accuracy of PET, CT and MRI Se (95% CI): PET 95% (90%-97%), CT 76% (70%-81%), MRI 78% (71%-84%) Sp (95% CI) : PET 90% (87%-93%), CT 59% (55%- 63%), MRI 76% (71%-80%) FDG-PET had significantly better sensitivity and specificity than CT and MRI. FDG-PET was the best modality for diagnosis of local residual or recurrent nasopharyngeal carcinoma	MEDLINE, EMBASE databases, CBMdisc databases for Chinese Articles, Sciencedirect, Springlink, Scopus, Cochrane Database of Systematic Review, and Database for Chinese Technological Journals January 1990 to May 2007 SPSS 13.0 for Windows (SPSS, Chicago, Ill), and Meta-DiSc Meta-regression analysis

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Brouwer 2008	30 consecutive patients suspected of recurrent laryngeal carcinoma after radiotherapy The suspicion was either raised by symptoms, such as voice deterioration, pain, dyspnea or dysphagia or by physical exam (i.e., office laryngoscopy).	**FDG-PET** Reference standard: Biopsy in the 12 months following FDG-PET	Diagnostic performance of PET to diagnose recurrence of laryngeal carcinoma after radiotherapy	8 patients had biopsy proven recurrent laryngeal carcinoma (27%). Se: 88% (95% CI 53–98%) Sp: 82% (95% CI 62–93%). *Conclusion*: FDG-PET is a promising technique to detect recurrent laryngeal carcinoma after radiotherapy and selecting patients for direct laryngoscopy.	FDG-PET scans were revised and assessed by 9 nuclear medicine physicians and residents in training, blinded for clinical results. Data show disparities among FDG-PET readers with potential impact on patient.

Monitoring of treatment response

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with head and neck cancer	**FDG-PET** Reference standard: histopathology from surgery or biopsy OR disease-specific survival and overall survival	Six studies (Dietz 2002, Kitagawa 2003a & 2003b, Kunkel 2003, McCollum 2004, Nam 2005), including 162 patients used PET to predict response to therapy (mainly after neoadjuvant therapy). These studies did not clearly demonstrate the value of PET, with a number of false classifications of response. No data about Se and Sp were provided.	See above

RT Planning

	Population	Index test	Results	Comments
HTA reports				
MSAC 2008	Patients with newly diagnosed head and neck cancer or unknown primary site	**FDG-PET** **FDG-PET/CT** Prior tests: Clinical evaluation (incl.	NCCHTA 2007 (Facey 2007) + 3 studies reporting the therapeutic impact of PET on treatment planning were identified (Connell 2007; Ha 2006; Wartski 2007). - Ha 2006: an overall change in management due to PET/CT was seen in 11 patients (31%); treatment plan	The quality of all three studies was limited by the fact that pre-PET plans were not reported, although it is stated in the methods section that they had been generated

	Population	Index test	Results	Comments
		CT, MRI) Reference standard: Panendoscopy (biopsy) or follow-up	alterations included addition of chemotherapy or radiotherapy so that multimodality approach was favoured. - Connell 2007: In 4/35 patients (11%), PET had a 'high impact' by changing the treatment modality. In 10/35 patients (29%) the impact was considered 'medium' as radiotherapy planning technique or dose was altered. - Wartski 2007: PET positive in 26/38 cases. In 23 patients, PET/CT provided additional information which altered the treatment plan The main change in management affects radiotherapy. Where this is based on more accurate staging of regional lymph nodes, this should result in an improvement in patient outcomes.	
NCCHTA 2007	Patients with head and neck cancer	**FDG-PET** Comparators: CT or MRI Reference: regions checked on dissection (Schwartz)	Three studies (Nishioka 2002, Scarfone 2004, Schwartz 2005) totalizing 47 patients used PET in RT planning, which resulted in change in gross tumour volume or the number of irradiated nodes in several patients. *Conclusion* : There are insufficient data to recommend the use of PET for RT planning in place of CT or MRI (differences in GTV and number of irradiated nodes are reported between methods)	See above
UK- NCCHTA 2007 (Facey et al. 2007)	Patients with head and neck cancer	**FDG-PET/CT** Reference standard: histopathology, additional imaging, clinical follow-up OR none Comparator: CWU	Three studies (Ciernik 2003, Koshy 2005, Paulino 2005) in a total of 88 patients used PET/CT for RT planning and showed changes (increase or decrease) in volume or dose compared with CT. *Conclusion* : There are insufficient data to recommend the use of PET/CT for RT planning in place of CT (differences in GTV and number of irradiated nodes are reported between methods)	See above

Prognosis

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Cicone 2008	35 successive patients with a clinical or iconographical suggestive SCC of the head and neck (for staging) and followed until death or until last follow-up (prognosis)	Pre-treatment **FDG-PET**	Disease-free survival (DFS) Overall survival (OS)	Follow-up: from diagnosis to 27 months (~) Median disease-free survival was 12.5 months (0-26.9 months). Univariate Cox regression: - only SUV mean bone marrow activity was predictive of DFS (p=0.05) and OS (p=0.028) Multivariate model (covariates: age and gender): - SUV mean bone marrow activity retained its prognostic value for DFS (p=0.04) as for OS (p=0.03) Kaplan-Meier analysis Dichotomized mean SUV of bone marrow : > 1.3 (better DFS) vs ≤ 1.3 (worse DFS) *Conclusion*: SUV mean bone marrow has a significant prognostic value in patients with head and neck cancer	This study also reported results for staging, but due to the impossibility to fill-in the 2*2 table, these results were not reported here. Overall survival was defined at the time from initial diagnosis until death or until last follow-up (right censored data). Overall survival took all deaths into account (13 deaths)

COLORECTAL CANCER

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with suspected colorectal cancer	**FDG-PET** Objective: Diagnosis of primary tumour	One HTA report identified (DACEHTA 2001), included in previous KCE report. One primary study found (Friedland 2005), including 45 patients with neoplastic colonic polyps. Reference standard: histopathology. PET had a sensitivity of 62% to detect malignant lesions, with specificity of 100%. It was noted that PET only detected one in six tumours that were less than 2 cm.	Good-quality HTA *Search date*: Aug 2005 *Databases*: Medline, EMBASE, Cochrane Library, HTA database, DARE, individual contacts through INAHTA No meta-analysis

Staging (primary colorectal cancer)

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with colorectal cancer (mixed primary and recurrent)	**FDG-PET** Objective: Staging (including evaluation of resectability of liver metastases)	One systematic review identified (Kinkel 2002), included in previous KCE report. Mixed population (primary and recurrent colorectal cancer). Reference standards: histopathology, core biopsy, cytology, or follow-up (min. 6 months). Two additional primary studies (Heriot 2004; Kantorova 2003 [included in previous KCE report]) evaluated PET before initial therapy for colorectal cancer. Heriot et al. showed that 8 out of 46 patients (17%) had their therapy altered as a result of PET. Five additional primary studies (Arulampalam 2004; Desai 2003; Rosa 2004 [included in previous KCE report]; Topal 2001; Truant 2005) reported staging in those considered eligible for resection of colorectal liver metastases (mixed populations: primary and recurrent colorectal cancer). Reference standards: histopathology and follow-up. PET was found to be more accurate than CT for detection of liver metastases (Se 100% vs. 47%) (Arulampalam 2004). Three studies showed that PET influenced or could have influenced therapy in 9%, 21% or 39% of patients. Two studies noted that 6% and 15% had staging incorrectly changed.	See above. Impossible to separate results for primary and recurrent colorectal cancer.
NCCHTA 2007	Patients with colorectal cancer	**FDG-PET/CT**	Two prospective primary studies identified (Francis 2003, Selzner 2004; mixed populations, in total 93 patients), showing that	See above

	Population	Index test	Results	Comments
			PET/CT had high sensitivity to detect primary tumours (100% in Francis 2003, per-lesion analysis) and liver metastases (91% in Selzner 2004). The largest study (Selzner 2004, n=76) showed that the specificity of PET/CT to detect liver metastases was 20% higher than CT. For local recurrence, both PET/CT and CT had high specificity, but PET/CT had 40% higher sensitivity than CT.	
Systematic reviews				
Bipat 2005	Patients with colorectal cancer and suspected liver metastasis (unclear if primary and/or recurrent colorectal cancer)	**FDG-PET** Comparators: CT MRI Objective: Detection of liver metastases	Twenty-one primary studies on FDG-PET included. Mixed reference standards, no further details specifically for PET studies. Sensitivity estimates on a per-patient basis for nonhelical CT, helical CT, 1.5-T MR imaging, and FDG-PET were 60.2%, 64.7%, 75.8%, and 94.6%, respectively; FDG-PET was the most accurate modality.	Good-quality SR *Search date*: December 2003 *Databases*: Medline, EMBASE, Cinahl, Cancerlit, SumSearch, Web of Science, CDSR *Languages*: English, French, German Meta-analysis performed

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET: staging (N- and/or M-staging)					
Llamas-Elvira 2007	Patients with a histological diagnosis of colorectal cancer (n=104)	**FDG-PET** Comparator: CT Standard: Histopathology (n=90) or clinical/imaging follow-up (n=14)	N-staging (n=90; partial verification: results not reported) M-staging (n=104)	M-staging: *Sensitivity (95%CI)*: PET 89% (64-98%), CT 44% (22-69%) *Specificity (95%CI)*: PET 93% (85-97%), CT 95% (88-98%) PET results led to modification of the therapy approach in 50% of patients with unresectable disease. PET revealed unknown disease in 19%, changing the staging in 13% and modifying the scope of surgery in 12% (with a change in the therapeutic approach in 18% of those patients with rectal cancer).	Prospective study. Consecutive patients. Differential verification.
FDG-PET/CT: N-staging					
Tateishi 2007	Patients with histologically proven	**FDG-PET/CT** Contrast-enhanced vs.	N-staging (n=53)	Pararectal nodes: *Sensitivity*: nePET/CT 73% (22/30;	Retrospective study.

	Population	Index test	Outcome	Results	Comments
	rectal cancer (n=53)	non-enhanced Standard: Histopathology (n=53)		95%CI 54-88%), cePET/CT 90% (27/30; 95%CI 73-98%) *Specificity*: nePET/CT 57% (13/23; 95%CI 34-77%), cePET/CT 78% (18/23; 95%CI 56-93%) Internal iliac nodes: *Sensitivity*: nePET/CT 60% (9/15; 32-84%), cePET/CT 73% (11/15; 45-92%) *Specificity*: nePET/CT 82% (31/38; 66-92%), cePET/CT 87% (33/38; 72-96%) Obturator nodes: *Sensitivity*: nePET/CT 50% (5/10; 19-81%), cePET/CT 80% (8/10; 44-97%) *Specificity*: nePET/CT 84% (36/43; 69-93%), cePET/CT 91% (39/43; 78-97%)	
FDG-PET/CT: evaluation of liver metastases					
Coenegrachts 2009	Patients with suspected colorectal liver metastases (n=24) (suspicion based on ultrasound or biochemistry)	**FDG-PET/CT** (contrast-enhanced) Comparator: MRI Standard: Intra-operative ultrasound with histology (n=18) or clinical/imaging follow-up (including MRI) (n=6)	Detection of liver metastasis (n=24)	Sensitivity: 96% (95%CI 79-100%) PPV: 100% (95%CI 85-100%)	Prospective study. Consecutive patients. Differential verification. Incorporation bias for MRI! Unclear if synchronic or metachronic liver metastases.
Kong 2008	Patients with colorectal cancer and known or suspicion of liver metastases (suspicion based on other imaging, including CT) (n=65)	**FDG-PET/CT** Comparators: MRI Contrast-enhanced CT (ceCT) Standard: Histopathology (n=23) or clinical/imaging follow-up	Detection of extrahepatic disease (no 2x2 table!) Detection of liver metastases	Extrahepatic disease: Change in management in 17%. Three false-positives. Detection of liver metastases: *Sensitivity*: PET 98% (91-100%), MRI 98% (91-100%) *Specificity*: PET 100% (40-100%), MRI 100% (40-	Retrospective study. Representative spectrum?? Cfr. mix of known and suspicious liver metastases. Differential verification. Unclear if blinded. No details on how follow-up was done. Unclear if synchronic or metachronic liver metastases.

	Population	Index test	Outcome	Results	Comments
		(median 13 months; n=42)		100%)	
Chua 2007	Patients with suspected (how?) metastatic disease (n=131), including 75 patients with colorectal cancer (mixed primary and recurrent)	**FDG-PET/CT** Comparator: Contrast-enhanced CT (ceCT) Standard: Histopathology Clinical/imaging follow-up (including ceCT and FDG-PET/CT)	Detection of liver metastasis (n=75)	**Discordant data in text and tables!** Detection of liver metastasis: *Sensitivity*: PET/CT 94% (85-98%), ceCT 91% (82-97%) *Specificity*: PET/CT 75% (35-97%), ceCT 25% (3-65%) PET/CT altered patient management over ceCT in 25% of patients.	Retrospective study. Differential verification (no data provided on amount of patients with histopathologic proof). Incorporation bias!

Detection and staging of recurrence

	Population	Index test	Results	Comments
HTA reports				
MSAC 2008	Patients with either apparently resectable and potentially curable pulmonary or hepatic metastasis or suspected locoregional recurrence of CRC	**FDG-PET** **FDG-PET/CT**	Six HTA reports identified between 1999-2006 (MDH 1999, MSAC 2001, ICES 2001, AETMIS 2001, NHS 2003, NCCHTA 2007). Inclusion of 2 primary studies on diagnostic accuracy published since NCCHTA 2007: - Liu 2005 (n=37): already included in Zhang 2009 (see below). Se 89%, SP 89%. - Amthauer 2006 (n=68): unable to calculate 2x2 table. Inclusion of 1 primary study on change in patient management (Scott 2006): - Detection of recurrence: clinician-reported change in treatment modality or intent due to PET 61/93 (65%, 95%CI: 56–75%). - Staging of solitary metastases: Clinician-reported change in treatment modality or intent due to PET 41/98 (42%, 95%CI: 33–52%).	Good-quality HTA. *Search date*: Dec 2006 *Databases*: Medline, Pre-Medline, EMBASE, Current Contents, Cochrane Library Controlled Clinical Trials Registry, databases of ongoing studies No meta-analysis
NCCHTA 2007	Patients with colorectal cancer (mixed primary and recurrent)	**FDG-PET**	Three systematic reviews identified (Kinkel 2002, Dietlein 2003, DACEHTA 2001), all included in previous KCE report. Mixed populations (primary and recurrent colorectal cancer). Reference standards: histopathology, core biopsy, cytology, or follow-up.	See above. Impossible to separate results for primary and recurrent colorectal cancer.

	Population	Index test	Results	Comments
			Five additional primary studies (Arulampalam 2004; Desai 2003; Rosa 2004 [included in previous KCE report]; Topal 2001; Truant 2005) reported staging in those considered eligible for resection of colorectal liver metastases (mixed populations: primary and recurrent colorectal cancer). Reference standards: histopathology and follow-up. Overall, PET was found to be more accurate than comparators for detection of liver metastases. Three studies showed that PET influenced or could have influenced therapy in 9%, 21% or 39% of patients. Two studies noted that 6% and 15% had staging incorrectly changed. Four additional primary studies (Fukunaga 2005; Montravers 2004; Langenhoff 2002; Selvaggi 2003) assessed colorectal cancer recurrence. Sensitivity: 85-100% Specificity: 83-100%	
NCCHTA 2007	Patients with colorectal cancer	**FDG-PET/CT**	Three retrospective primary studies of PET/CT vs. PET identified (Cohade 2003, Even-Sapir 2004, Kim 2004). The 3 studies involved 157 patients. One study showed that both assessed recurrence accurately, while another showed slightly better sensitivity of PET/CT (96% vs. 88%) and higher specificity (89% vs. 74%). In the other trial 88% of patients were correctly staged with PET/CT vs. 71% of patients with PET.	See above
Systematic reviews				
Zhang 2009	Patients with suspected recurrent colorectal cancer	**FDG-PET** Objective: Detection (+ staging) of recurrent disease	Twenty-seven primary studies identified. Most studies used acceptable reference tests, including histopathologic examination, percutaneous biopsy specimens and serial CT scans and follow-ups (no further details provided). Distant metastasis or whole body involvement: Pooled sensitivity: 0.91 (95%CI 0.88–0.92) Pooled specificity: 0.83 (95%CI 0.79–0.87) Hepatic metastasis: Pooled sensitivity: 0.97 (95%CI 0.95–0.98) Pooled specificity: 0.98 (95%CI 0.97–0.99)	Good-quality SR *Search date*: January 2008 *Databases*: Medline, EMBASE Inclusion of English literature only *Meta-analysis* with Meta-DiSc v1.4 and STAT v9

	Population	Index test	Results	Comments
			Pelvic metastasis or local regional recurrence: Pooled sensitivity: 0.94 (95%CI 0.91–0.97) Pooled specificity: 0.94 (95%CI 0.92–0.96)	
Bipat 2005	Patients with colorectal cancer and suspected liver metastasis	**FDG-PET** CT MRI	Twenty-one primary studies on FDG-PET included. No details on reference standards for the PET studies. Sensitivity estimates on a per-patient basis for nonhelical CT, helical CT, 1.5-T MR imaging, and FDG-PET were 60.2%, 64.7%, 75.8%, and 94.6%, respectively; FDG-PET was the most accurate modality.	Good-quality SR *Search date*: December 2003 *Databases*: Medline, EMBASE, Cinahl, Cancerlit, SumSearch, Web of Science, CDSR *Languages*: English, French, German Meta-analysis performed
Wiering 2005	Patients with recurrent colorectal cancer undergoing hepatic resection for colorectal metastases	**FDG-PET** Objective: Evaluation of liver metastases	Thirty-two primary studies included. No details on used reference standards. Conclusions of authors: Despite apparent omissions in the literature, the combined sensitivity and specificity of FDG-PET clearly indicated that FDG-PET has added value in the diagnostic workup of patients with colorectal liver metastases. FDG-PET can be considered a useful tool in preoperative staging and produced superior results compared with conventional diagnostic modalities, especially for excluding or detecting extrahepatic disease.	Moderate-quality SR *Search date*: January 2004 *Databases*: Medline, EMBASE Meta-analysis performed, but few details available

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Sobhani 2008	Patients treated with curative R0 surgery for colon or rectal cancer (n=130)	Follow-up with **FDG-PET** at 9 and 15 months after surgery (in addition to conventional follow-up) (n=65) Comparator: Conventional follow-up (Conv), comprising 6 visits (physical examination, serum CEA and/or CA19-9, US every 3 months [except after 9	Overall rate of recurrence at 15 months Time-to-recurrence Time to second-line surgical intervention and/or drug treatment	Number of recurrences: PET n=25, Conv n= 21; p=0.50 Time-to-recurrence: No significant difference with regard to actuarial curves of recurrence (log-rank p=0.55). For all the patients with a recurrence, the time from baseline until detection of the recurrence was significantly shorter in the PET group than in the Conv group (12.1±3.6 vs. 15.4±4.9 months; p=0.01).	RCT: no information on randomisation procedure. Interpretation of PET without knowledge of CT results; no other information on blinding. Differential verification. Intention-to-treat analysis.

	Population	Index test	Outcome	Results	Comments
		and 15 months], chest X-ray every 6 months, and abdominal CT scans after 9 and 15 months) (n=65) Standard: Recurrence was identified from histological samples (biopsy or surgery) in all cases except in those with evidence of recurrence consisting of disseminated metastases or those for whom clinical examination, tumour markers and imaging procedures (routinely discussed during a multidisciplinary staff meeting) yielded consistently positive results	Overall rate of curative surgery	Impact on management: *Time-to-treatment*: PET 14.8±4.1 months, Conv 17.5±6 months; p=0.09 *R0 curative resections*: PET n=10, Conv n=2; p<0.01 No significant difference in mortality. Diagnostic accuracy: (per-protocol-analysis) Sensitivity 92% (22/24) Specificity 86% (39/43)	
Votrubova 2006	Patients with colorectal cancer after operation (n=84)	**FDG-PET** **FDG-PET/CT** Contrast-enhanced (n=30), non-enhanced (n=54) Standard: Histopathology or follow-up	Detection of intra-abdominal extrahepatic recurrence and extra-abdominal and/or hepatic recurrence (n=84)	Recurrence in general: *Sensitivity*: PET/CT 89% (76-96%), PET 80% (65-90%) *Specificity*: PET/CT 92% (79-98%), PET 69% (52-83%) Intra-abdominal extrahepatic recurrence: *Sensitivity*: PET/CT 88% (72-97%), PET 82% (65-93%) *Specificity*: PET/CT 94% (84-99%), PET 88% (76-96%)	Retrospective study. No exact details on how follow-up after PET/CT was done. Differential verification (no data provided on amount of patients with histopathologic proof). **Identified by Zhang 2009, but excluded due to using PET/CT.**

	Population	Index test	Outcome	Results	Comments
				Extra-abdominal and/or hepatic recurrence: *Sensitivity*: PET/CT 95% (74-100%), PET 74% (49-91%) *Specificity*: PET/CT 100% (94-100%), PET 88% (77-95%)	
FDG-PET/CT					
Bellomi 2007	Patients treated with radical surgery for rectal cancer who underwent FDG-PET/CT for a suspicion of local or distant recurrence (clinical follow-up or MDCT) (n=67)	**FDG-PET/CT** Comparator: MDCT Standard: Histology of biopsy or surgical specimens, or follow-up	Detection of local recurrence (n=67) Detection of hepatic recurrence (n=67) Detection of pulmonary recurrence (no full 2x2 table!)	Local recurrence: *Sensitivity*: PET/CT 93% (68-100%), MDCT 100% (78-100%) *Specificity*: PET/CT 98% (90-100%), MDCT 98% (90-100%) Hepatic recurrence: *Sensitivity*: PET/CT 100% (80-100%), MDCT 100% (80-100%) *Specificity*: PET/CT 100% (93-100%), MDCT 100% (93-100%) Pulmonary recurrence: *Sensitivity*: PET/CT 75%, MDCT 100%	Retrospective study. No exact details on how follow-up after PET/CT was done. Differential verification (no data provided on amount of patients with histopathologic proof).
Votrubova 2006	See above in this table.				

Monitoring of treatment response

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with colorectal cancer	**FDG-PET**	Six primary studies found (Anthauer 2004, Calvo 2004, Capirci 2004, Denecke 2005, Dimitrakopoulou-Strauss 2003, Guillem 2000). Mixed populations (5 studies including patients with locally-advanced rectal cancer, 1 study including patients with metastatic colorectal cancer). Mixed objectives (5 studies evaluating response to treatment [4 studies with neoadjuvant CRT, 1 study with second-line palliative chemotherapy]). Most studies used histopathology as reference standard (1 study used clinical response). In three studies 2x2 table calculable (Amthauer 2004, Capirci 2004, Denecke 2005): Se 79-100%, Sp 45-86%.	See above

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Cascini 2006	Patients with biopsy-proven locally-advanced rectal adenocarcinoma scheduled to undergo TME 8-10 weeks after completing CRT (n=33)	**FDG-PET** Before and 12 days after the beginning of the neoadjuvant CRT Standard: Histopathology (tumour regression grade [TRG])	Treatment response (n=33)	Significant correlation between TRG and early SUV changes ($p<0.0001$). Responders were identified correctly by an early decrease of the mean SUV of ≥52%. Optimum threshold value for the SUV-max = -42%: sensitivity 100%, specificity 87%.	Prospective study. Consecutive patients.
FDG-PET/CT					
Rosenberg 2009	Patients with uT3 rectal carcinoma scheduled to undergo surgery during week 5 after completion of CRT (n=30)	**FDG-PET/CT** Contrast-enhanced Before, 14 days after the beginning of the neoadjuvant CRT, and 4 weeks after completion Standard: Histopathologic tumour regression according to	Treatment response (n=29)	The mean (±SD) reduction of tumour FDG uptake in histopathologically responding compared to non-responding tumours was ☐44.3% (±20.1%) vs. ☐29.6% (±13.1%) ($p=0.085$) at day 14 and ☐66.0% (±20.3%) vs. ☐48.3% (±23.4%) ($p=0.040$) after completion of CRT.	Prospective study. One patient refused surgery.

	Population	Index test	Outcome	Results	Comments
		the classification of Becker		Optimum threshold value for reduction in FDG uptake at 14 days = -35%: sensitivity 74% (95%CI 48-91%), specificity 70% (95%CI 34-94%). Optimum threshold value for reduction in FDG uptake after completion = -57.5%: sensitivity 79% (95%CI 54-94%), specificity 70% (95%CI 34-94%).	
Capirci 2007	Patients with histologically proven rectal adenocarcinoma scheduled to undergo surgery 8–10 weeks after the end of CRT (n=48)	**FDG-PET/CT** Before and 4–6 weeks after completion of the neoadjuvant CRT Standard: Histopathology (tumour regression grade [TRG])	Treatment response (n=45)	No significant difference in either the ☐SUV or the response index (RI) values within each of the four TRG levels. Threshold value for SUV-max value = -66.2%: sensitivity 81%, specificity 79% (area under the curve=0.856, 95%CI 0.719–0.942, p<0.0001).	Prospective study. Consecutive patients. Three patients excluded from analysis (1 died before surgery, 2 patients still due for surgery)

Radiotherapy planning

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with rectal cancer	**FDG-PET**	One primary study found (Ciernik 2005), involving 11 patients. Comparator: CT. PET and CT produced similar RT planning regions.	See above

Prognosis

	Population	Index test	Results	Comments
HTA reports				
MSAC 2008	Patients with either apparently resectable and potentially curable pulmonary or hepatic metastasis or suspected locoregional recurrence of CRC	**FDG-PET** **FDG-PET/CT**	One ongoing RCT was identified (Oyen 2007). Preliminary results indicated that the 9-month disease-free survival of patients selected for surgery on the basis of PET was 66% (n=44) vs. 45% in the group selected for surgery without PET (n=49). 9-month disease-free survival of patients proceeding to hepatic resection did not significantly differ in the PET vs. the non-PET arm (72% versus 55%, p=0.14). However, the power of the study and/or the follow-up time may be too limited in those preliminary analyses to detect a significant difference.	See above.

	Population	Index test	Results	Comments
			One additional primary study (Scott 2006) demonstrated that patients with PET-detected disease not apparent on prior imaging have a higher risk of disease progression at 12 months than those without PET-detected extra sites of disease for both colorectal indications (suspected locoregional recurrence, RR 1.67, 95%CI 1.06–2.62; isolated metastases, RR 1.68, 95%CI 1.12–2.52).	

Primary studies

	Population	Index test	Outcome	Results	Comments
de Geus-Oei 2008	Patients with metastatic colorectal cancer who were scheduled to undergo palliative chemotherapy (n=61)	**FDG-PET** Before and at 2 (n=50) and 6 months (n=19) after the start of treatment Standard: Clinical/imaging follow-up (in case of inconclusive findings: US and/or additional PET)	Overall survival (Kaplan-Meier) Progression-free survival (Kaplan-Meier)	Increase in rates of death (p=0.017 for ΔSUV PET1–2; p=0.048 for ΔSUV PET1–3) and progression (p=0.035 for ΔSUV PET1–2; p=0.081 for ΔSUV PET1–3) associated with worse response as assessed by PET on Cox proportional regression analysis. The OS and PFS analysis showed a significant predictive value at broad ranges of ΔSUV cut-off levels.	Prospective study. Consecutive patients. Overall median follow-up: 18.5 months (range 5-45 months).
Nakagawa 2008	Patients with primary rectal cancer treated with preoperative radiotherapy (n=59)	**FDG-PET** Before treatment and 2-3 weeks after radiotherapy	Overall survival (Kaplan-Meier)	In multivariate analysis, residual tumour and SUV after radiotherapy were significant prognostic factors for survival. Median survival and 5-year overall survival comparing SUV<5 vs. >5 were 95 vs. 42 months and 70 vs. 44 percent, respectively (p=0.042).	Prospective study. No exact information on duration of follow-up.
Capirci 2006	Patients with histologically proven rectal cancer undergoing neoadjuvant CRT (n=88)	**FDG-PET** Seven weeks after completion of CRT	Disease-free survival Overall Survival	Multivariate analysis showed that pathologic stage and FDG PET findings were independent prognostic predictors of both overall survival and disease-free survival. 5-year overall survival was 91% in patients with a negative PET after CRT vs. 72% in those with a positive PET (p = 0.024) after CRT, whereas disease-free survival was 81% and 62% (p =	Prospective study. Consecutive patients. All patients underwent surgery. No exact information on duration of follow-up.

	Population	Index test	Outcome	Results	Comments
				0.003) for those with the negative and positive PET findings, respectively.	
Kalff 2006	Patients with biopsy-proven rectal cancer considered suitable for aggressive neoadjuvant chemoradiation (n=34)	**FDG-PET** At staging and after completion of chemoradiation (1 patient had a mid-therapy PET)	Overall survival (Kaplan-Meier) Time-to-progression	PET response was highly significantly associated with overall survival ($p<0.0001$) and time to progression ($p<0.0001$). The percentage of maximum SUV change after chemoradiation was not predictive of survival in partial metabolic response patients.	Prospective study. Consecutive patients. Median follow-up: 3.1 years (range 1.6-4.7 years). Partial verification for diagnostic accuracy for evaluation of treatment response (results not presented).

MELANOMA

Staging

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with primary and recurrent malignant melanoma	**FDG-PET** Objective: Staging	Two systematic reviews identified (Mijnhout 2001, DACEHTA 2001), both mixing studies with primary and recurrent malignant melanoma. Both SRs were included in the previous KCE report. Twelve additional studies identified on staging alone. Mixed reference standards (majority histopathology and/or SLNB). Nine of these studies showed highly consistent results that PET had poor sensitivity (generally <20%) to detect *regional lymph-node activity* in early-stage patients (n=528; Se 0-40%): Wagner 2005 (n=144): Se 21% (10-36%), Sp 97% (93-99%) Fink 2004 (n=48) : Se 13% (2-47%), Sp 100% (NB : already included in previous KCE report) Hafner 2004 (n=100) : Se 8% (1-25%), Sp 100% (95-100%) Havenga 2003 (n=55) : Se 15% (NB : already included in previous KCE report) Longo 2003 (n=25) : Se 22% Belhocine 2002 (n=21) : Se 14% (0-28%), Sp 93% (83-103%) Reinhardt 2002 (n=67) : Se 92% (for LN and distant metastases) Acland 2001 (n=50): Se 0% (95%CI 0-23%) Kokoska 2001 (n=18) : Se 40% For *distant metastases*, there were several false positives and one study in which	Good-quality HTA *Search date*: Aug 2005 *Databases*: Medline, EMBASE, Cochrane Library, HTA database, DARE, individual contacts through INAHTA No meta-analysis

	Population	Index test	Results	Comments
			the sensitivity was only 4% (Wagner 2005). PET sensitivity varied between 40 and 100% in the three primary studies in later stage disease (Ghanem 2005, Gulec 2005, Vereecken 2005 [included in previous KCE report]). Again, sensitivity in small lesions was poor (13% in Gulec 2003). For later stage disease, comparative results are varied. In one study PET was less sensitive than MRI, but in another PET was superior to CT/MRI and led to more changes in treatment. Three additional primary studies in mixed primary and recurrent populations found (Finkelstein 2004, Jenicke 2001, Kurli 2005). Mixed reference standards.	
Systematic reviews				
Krug 2008	Patients with histopathologically proven cutaneous malignant melanoma (CMM)	**FDG-PET(/CT)** Standard: Histopathology or clinical-radiological follow-up Objective: Initial staging	Twenty-eight studies included (2905 patients, 2096 underwent PET and 809 PET-CT). Pooled estimates of FDG-PET for the detection of metastasis in the initial staging of CMM were: *Sensitivity* 83% (95%CI 81%-84%) *Specificity* 85% (95%CI 83%-87%) *Positive likelihood ratio* (LR) 4.56 (95%CI 3.12-6.64) *Negative LR* 0.27 (95%CI 0.18-0.40) *Diagnostic odds ratio* (OR) 19.8 (95%CI 10.8-36.4) Early stages (n=755): *OR* 4.3 (95%CI 1-18) *Sensitivity* 60% (95%CI 54%-60%) Advanced stages (n= 2150): *Sensitivity* 86% (95%CI 84%-87%) *Specificity* 87% (95%CI 85%-88%) Eight studies suggested that FDG-PET was associated with 33% disease management changes (range 15%–64%).	Good-quality SR *Search date*: March 2007 Databases: Medline, Embase, Web of Science, Cochrane Database of Systematic Reviews *Meta-analysis* using random-effects symmetric summary receiver operating characteristic (SROC) curve analysis. Software: Meta-Disc, version 1.4

Primary studies

	Population	Index test	Outcome	Results	Comments
Singh 2008	Patients with histopathologically proven melanoma (stage I/II) (n=52)	**FDG-PET/CT** Standard: Histopathology (SLNB)	N-staging	*Sensitivity*: 14% (95%CI 3-44%) *Specificity*: 95% (81-99%)	Prospective study.
Akcali 2007	Patients with clinically evident stage III or IV cutaneous melanoma (n=39)	**FDG-PET/CT** Standard: Biopsy	M-staging	*Sensitivity*: 92% (11/12; 95%CI 62-100%) *Specificity*: 92% (24/26; 95%CI 75-99%)	Retrospective study. Exclusion of 1 patient (reason unclear). No information on blinding.
Strobel 2007a	Patients with high-risk melanoma (n=124)	**FDG-PET/CT** Standard: Histological examination, other imaging modalities (such as MRI and/or PET/CT follow-up), and/or clinical follow-up for a minimum of 6 months	M-staging	*Sensitivity*: PET/CT: 85% (45/53; 95%CI 72-93%) PET/CT with dedicated CT interpretation: 98% (52/53; 95%CI 90-100%) *Specificity*: PET/CT: 96% (68/71; 95%CI 88-99%) PET/CT with dedicated CT interpretation: 94% (67/71; 95%CI 86-98%)	Prospective study. Consecutive patients. Differential verification. Incorporation bias.
Strobel 2007b	Patients with high-risk melanoma and elevated S-100B levels (n=47)	**FDG-PET/CT** (contrast-enhanced) Comparator: S-100B Standard: Composite reference standard (cytological, histological, MRI and PET/CT follow-up findings as well as clinical and S-100B follow-up)	N-staging M-staging	Lymph nodes: *Sensitivity* 89% (8/9; 95%CI 52-100%) *Specificity* 100% (38/38; 95%CI 91-100%) Distant metastasis: *Sensitivity* and *specificity* 100%	Retrospective study. Consecutive patients. Differential verification. Incorporation bias. No information on blinding.

Detection of recurrence

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with suspected recurrence of melanoma	**FDG-PET** Objective: Detection of recurrence	Two primary studies identified (Mijnhout 2002, Stas 2002 [lesion-based analysis]) involving a total of 152 patients. Stas demonstrated that PET had sensitivity and specificity of at least 85% for all tumour sites combined and for the individual areas of lung and lymph nodes (reference standard: histopathology/ follow-up/none). PET accuracy was poorer in skin and brain metastases. The diagnostic accuracy study also reported that PET affected 30% of patients' therapy. The other study (Mijnhout 2002) was a well-designed patient management study, which found that PET contributed to change in therapy in at least 34% of patients.	See above.

Staging of recurrence

	Population	Index test	Results	Comments
HTA reports				
MSAC 2008	Patients with biopsy-proven recurrence of malignant melanoma	**FDG-PET** **FDG-PET/CT** Objective: Staging of recurrent disease	Six HTA reports identified between 1999 – 2006 (MDH 1999, MSAC 2001, ICES 2001, AETMIS 2001, NHS 2003, NCCHTA 2007). One additional SR identified (Ratko 2006). Earlier search dates than NCCHTA 2007. One additional primary study on diagnostic accuracy identified (Reinhardt 2006). One additional primary study on change in management identified (Fulham 2006). Reinhardt 2006 (n=65): N-staging: PET Se 100% (98-100%), Sp 91% (84-98%); PET/CT Se 100% (98-100%), Sp 100% (98-100%). Significantly better sensitivity than CT alone. M-staging: PET Se 98% (94-100%), Sp 88% (80-96%); PET/CT Se 100% (98-100%), Sp 96% (91-100%). Significantly better sensitivity for PET/CT compared to CT alone. Fulham 2006 (n=134): 10.4% change in surgical procedure, 13.4% change from surgery to chemotherapy.	Good-quality HTA *Search date*: Dec 2006 *Databases*: Medline, EMBASE, Pre-Medline, Cochrane Library Controlled Clinical Trials Registry, Current Contents No meta-analysis
NCCHTA 2007		**FDG-PET** Objective: Staging of recurrent disease	See above in table 'Staging'.	See above.

BREAST CANCER

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients who have an abnormal mammogram or palpable breast mass and have been referred for breast biopsy	**FDG-PET** Reference standards: cytological aspiration and histopathology	One systematic review identified (AHRQ 2001): already included in previous KCE report. Additional primary study (Heinisch 2003) compared PET and MRI in 36 women with suspicious lesions on mammography or clinical examination. PET Se 76% (95% CI: 52% - 91%) Sp 73% (95% CI: 45% - 91%) MRI Se 95% (95% CI: 74% - 99%) Sp 73% (95% CI: 45% - 91%)	Good-quality HTA *Search date*: Aug 2005 *Databases*: Medline, EMBASE, Cochrane Library, HTA database, DARE, individual contacts through INAHTA Meta-analysis using random-effects Trials only include patients with suspicious mammograms or palpable masses, so prevalence is high and mean tumour size was large. Hence, report states that evidence is required in other patients
AHRQ 2006	Patients who have suspicious breast lesions (abnormal mammogram and/or physical examination and/or ultrasound examination)	**FDG-PET** Comparators: MRI, US, scintimammography Reference standard: biopsy	*Objective*: to determine if available non invasive diagnostic test (PET/MR/US/scintimammography) are sufficiently accurate to exclude malignancy, avoiding women with an abnormal mammogram to perform biopsy. 69 publications were included: - 9 of 18-FDG PET scanning (8 WBS, 1 gamma camera). - 45 of scintimammography (SCM) - 19 of MRI - 8 of ultrasound Some publications reported data for more than one technology *For suspicious lesions* Se: PET (82.2%); MRI (92.5%); US (86.1%) Sp: PET (78.3%); MRI (72.4%); US (66.4%) *For non palpable lesions* Se: SCM (68.7%)	High quality HTA Search date : April 2005 Databases: PubMed, EMBASE, Clinical Trials, Cochrane Databases, ECRI databases, CRISP, Controlled Trials, Database of Abstracts of Reviews of Effectiveness (DARE), U.S. Centers for Medicare & Medicaid Services. The quality of all of the studies was moderate.

	Population	Index test	Results	Comments
			Sp: SCM (84.8%) In USA, after an abnormal mammogram, women have a level of risk of cancer = 20%. All technologies could reduce the need for biopsy (a) but each would miss some cancers (b). At this average risk level, in 1 000 women with: - a negative PET scan, 924 (a) but 76 (b) - a negative SCM, 907 (a) but 93 (b) - a negative MRI, 962 (a) but 38 (b) - a negative US, 950 (a) but 50 (b) Future studies could overturn these findings. *Conclusion*: MRI is a more valuable tool than PET to give a diagnosis (higher sensitivity and higher NPV). However, if a less than 2% risk of having breast cancer with a negative diagnostic test is considered an acceptable level of risk for a diagnostic test to reliably preclude biopsy, none of these tests was sufficiently accurate to replace biopsy for women at average risk of breast cancer. For non palpable lesions, data were insufficient to estimate the accuracy of PET, MRI or US. SCM was not sufficiently accurate to avoid biopsy. For palpable lesions, data were insufficient to estimate the accuracy of PET, MRI, US and SCM.	
Systematic reviews				
Bourguet 2006	Patients with suspicion of breast cancer	**FDG-PET**	No change since 2003. *Standard*: PET is not indicated in the diagnosis of breast cancer (evidence level A).	Update of a previous systematic review (2003) Literature search in Medline (2003-November 2005) + OVID alerts Language restrictions: French and English

Staging: axillary lymph nodes

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Extent of tumour in ALN in patients with confirmed primary breast malignancy, no palpable ALN metastases (cN0) and no evidence of distant metastases	**FDG-PET** Reference standards: ALND ALND + SNB	One systematic review (BCBSA 2003) already included in previous KCE report, and four additional primary studies (Fehr 2004, Lovrics 2004, Wahl 2004, Zornoza 2004). ALND as ref.: PET Se = 40–93% PET Sp = 87–100% ALND + SNB as ref.: PET Se = 20–50% PET Sp = 82–100% Prevalence of node-positive disease = 33–64%, so 36–67% patients with PET negative would have axillary disease undetected if further tests were not undertaken. *Conclusion*: PET cannot be used to avoid ALND in patients with clinically N0 axillae, because of unacceptably low sensitivity. With this level of false negatives, if patients did not go on to have standard diagnostic tests, modelling suggests that under-treatment would be associated with absolute difference in 10-year survival of 8.2%. *Recommendation*: PET cannot be reliably used to avoid ALND.	See above
Systematic reviews				
Sloka 2007	Patients with breast cancer	**FDG-PET** Reference standards: Histology via ALND / SNB / histology / histology + ALND / SNB +histo via ALND	**19 studies for staging axillary lymph nodes were considered in this systematic review.** In 3 high-quality studies (of which 2 were already included in previous KCE report: Wahl 2004, Zornoza 2004), i.e. studies with broad generalizability to a variety of patients and no significant flaws in research methods (Wahl 2004, Zornoza 2004, Greco 2001): - sensitivity : 61 – 94% - specificity : 80 – 98%	Literature search in December 2005 (MEDLINE, Current Contents and EMBASE) restricted to English, Spanish and French language articles. Due to the high heterogeneity between studies, meta-analysis was not performed.

	Population	Index test	Results	Comments
			Recommendation: Authors recommend that further studies be performed that control for contributory variables (patient position, etc) in order to explain the variability of study results. Avoid older studies (< 1992) due to the increased accuracy of new scanners.	
Bourguet 2006	Patients with breast cancer	**FDG-PET**	**1 primary study (Zornoza 2004):** already included in previous KCE report. No change since 2003: PET is unable to detect microscopic lymph node metastasis. *Option*: PET enables documentation of loco-regional invasion and metastatic spread in the initial staging of invasive breast cancer (evidence level B2). *Recommendation*: the place of PET in the initial staging of invasive breast cancer remains to be established.	Update of a previous systematic review (2003) Literature search in Medline (2003-November 2005) + OVID alerts Language restrictions: French and English

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Ueda 2008	183 patients having primary breast cancer proven by core needle biopsy who are operable	**FDG-PET/CT** Comparator: axillary US Standard reference: ALND and/or SNB	Diagnostic performance of PET/CT and AUS in assessing axillary status: Se and Sp	18-FDG PET/CT - visual assessment: Se: 58% (95% CI: 44% - 70%) Sp: 95% (95% CI: 89% - 98%) - SUV cut-off point 1.8 Se: 36% (95% CI: 24% - 49%) Sp: 100% (95% CI: 96% - 100%) AUS Se: 54% (95% CI: 31% - 55%) Sp: 99% (95% CI: 95% - 100%) Visual assessment of 18F-FDG uptake combined with AUS Se: 64% (95% CI: 51% - 76%) Sp: 94% (95% CI: 88% - 97%)	Prospective study Possibility of review bias: unclear

	Population	Index test	Outcome	Results	Comments
				Conclusion: performance of 18F-FDG PET/CT was almost equivalent to that of AUS for detecting of ALN involvement in patients with primary breast cancer. Sensitivity was low in both cases. The combination of these 2 exams slightly increased sensitivity. When it is difficult to judge the axillary staging using AUS alone, metabolic approach of 18F-FDG PET/CT for axillary staging would enable a much more confident diagnosis.	
Veronesi 2007	236 patients with breast cancer and clinically negative axilla	**FDG-PET** Comparator: SNB Standard reference: ALND	Diagnostic performance of PET and SNB in assessing axillary status: Se and Sp	103 out of the 236 patients (44%) had metastases in axillary nodes 18 FDG-PET: Se: 37% (95% CI: 28% - 47%) Sp: 96% (95% CI: 91% - 99%) SNB: Se: 96% (95% CI: 90% - 99%) Sp: 100% (95% CI: 96% - 100%) Conclusion: The high specificity of PET indicates that patients who have a PET-positive axilla should perform an ALND rather than an SNB for axillary staging. In contrast, when FDG-PET is negative at the axilla, its reliability is very low and axillary SNB becomes imperative.	Prospective study conducted from September 2003 to April 2005 in Italy
Gil-Rendo 2006	150 women with breast cancer: histologically proven carcinoma of the breast with clinically and ultrasonographically non-suspicious axillary	**FDG-PET** Standard reference: ALND	Diagnostic performance of PET in assessing axillary status: Se and Sp	In the first group of 150 women who had preoperative PET and ALND, the sensitivity and specificity for detecting axillary status were: Se: 90% (95% CI: 83% - 97%) Sp: 98% (95% CI: 93% - 99%) PET detected axillary involvement in 64	Prospective study on 275 women (2 subgroups). In a first group (150 women), ALND was performed regardless of PET results with the aim of evaluating the Se and Sp of the technique. In a second

	Population	Index test	Outcome	Results	Comments
	lymph nodes, eligible for primary treatment by breast conservation or mastectomy			of 71 patients (7 false negatives) and correctly diagnosed 78 of 79 patients without axillary metastases. Conclusion: The high sensitivity and the high specificity of PET suggest that FDG uptake in the axilla could be an indication for full ALND without previous SLNB	group (125 women), the axillary examination was complemented by SLNB only in those with no pathological axillary uptake on the FDG-PET scan.
Kumar 2006	80 women with a histological diagnosis of breast cancer and clinically negative axillary nodes	**FDG-PET** Standard reference: SLNB or ALND	Diagnostic performance of PET in assessing axillary status: Se and Sp	36 out of the 80 patients (45%) had metastases in axillary nodes 18 FDG-PET: Se: 44% (95% CI: 28% - 62%) Sp: 95% (95% CI: 83% - 99%) Conclusion: FDG PET cannot replace histological staging using SLNB in patients with breast cancer. The high specificity of PET indicates that patients who have a PET-positive axilla should perform an ALND rather than an SLNB for axillary staging. In contrast, FDG-PET showed poor sensitivity in the detection of axillary metastases, confirming the need for SLNB in cases where PET is negative in the axilla.	Prospective study in USA

Staging: metastases

Systematic review				
Shie 2008	Female patients with breast cancer, of all ages in any disease stage regardless of treatment status	**FDG-PET** Comparator: Bone scintigraphy Reference standards: CT, MRI or bone biopsy with clinical follow-up longer than	**Six articles** comparing FDG-PET and bone scintigraphy for the detection of osseous metastasis from breast cancer (Ohta 2001, Nakai 2005, Abe 2005, Gallowitsch 2003, Uematsu 2005, Yang 2002). The first 3 studies presented patients-based data whereas the 3 other studies reported lesions-based data. FDG-PET: The pooled patient-based sensitivity was 81% (95%	SR: MEDLINE, CINAHL, and EBM Review databases from January 1995 to November 2006. Results from meta-analyses were not considered here due to a lack of methodology description and results

		6 months	CI: 70%–89%), specificity was 93% (95% CI: 84%–97%). Bone scan: The pooled sensitivity was 78% (95% CI: 67%–86%), specificity was 79% (95% CI: 40%–95%). For the 3 patients-based studies, the differences in sensitivity and specificity of FDG-PET and bone scintigraphy in diagnosing osseous metastases are not statistically significant. *Conclusion:* It remains inconclusive whether FDG-PET or bone scintigraphy is superior in detecting osseous metastasis from breast cancer. FDG-PET does have a higher specificity and may better serve as a confirmatory test than bone scintigraphy.	reported
Bourguet 2006	Patients with breast cancer	**FDG-PET**	No change since 2003 *Recommendation:* FDG-PET allows work-up for detecting metastatic breast cancer in patients clinically suspected of metastasis (level of evidence B).	Update of a previous systematic review (2003) Literature search in Medline (2003-November 2005) + OVID alerts Language restrictions: French and English

Primary studies

	Population	**Index test**	**Outcome**	**Results**	**Comments**
FDG-PET					
Nakai 2005	89 breast cancer patients	**FDG-PET** Comparator: Bone scintigraphy Reference standard: bone biopsy / multi-slice CT for visual classification	Performance of PET/bone scintigraphy in diagnosing bone metastases: Se and Sp	Bone scintigraphy: Se: 78% (95% CI: 64% - 88%) Sp: 82% (95% CI: 65% - 92%) FDG-PET: Se: 80% (95% CI: 66% - 89%) Sp: 88% (95% CI: 71% - 96%) According to the CT image type, the visualisation rate of bone scintigraphy / FDG-PET was 100%/55.6% for the blastic	Retrospective study Selection criteria for patients were not described. Possibility of disease progression and review bias (unclear).

	Population	Index test	Outcome	Results	Comments
				type, 70.0%/100.0% for the lytic type, 84.2%/94.7% for the mixed type and 25.0%/87.5% for the invisible type. The visualisation rates of bone scintigraphy for the blastic type and FDG-PET for the invisible type were significantly higher. Conclusion: FDG-PET is useful in detecting bone metastases from breast cancer. It does, however, suffer from the drawback of a lower visualisation rate for osteoblastic bone metastases. FDG-PET should not be used alone to search for bone metastases in breast cancer patients.	

Restaging

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Detection of distant metastasis in patients with diagnosis of breast cancer	**FDG-PET**	**One meta-analysis (Isasi 2005)** already included in previous KCE report. **One small study (Uematsu 2005)** also included in previous KCE report. Per-lesion-analysis.	See above

Detection of recurrence

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with breast cancer and clinical suspicion of recurrence (with arm pain or other symptoms referable to the brachial plexus)	**FDG-PET** Reference standard: histopathology/follow-up	**One systematic review (BCBS 2003)** and **one additional primary study (Goerres 2003)** both included in previous KCE report.	See above

Monitoring of treatment response

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with locally advanced breast cancer	**FDG-PET** Reference standard: various	**One systematic review (Krak 2004)** including 8 primary studies, and **one additional primary study (Kim 2004)** in 50 women showed that midcourse PET scans could predict response to neoadjuvant chemotherapy in locally advanced breast cancer. **Kim 2004** (Reference Standard: pathological response from surgery): 8% of the patients had pathological complete response (CR) and 46% had pathological partial response (PR). Ten per cent of patients had clinical CR and 52% had clinical PR. For a reduction rate in SUV=79%, Se and Sp are respectively 85% and 83%. For a reduction rate in SUV=88%, Se and Sp are respectively 100% and 56%. This study suggests a possible predictive value of FDG PET for the assessment of the pathological response of primary breast cancer after neo-adjuvant chemotherapy. However, these findings deserve further investigation on a larger number of patients. There is no reliable evidence that PET scanning can predict response in axillary lymph nodes, or that post-treatment scans are able to detect microscopic residual foci of disease. **Two small primary studies (Gennari 2000, n=9; Schwarz 2005, n=11)** used PET midcycle to predict clinical response to chemotherapy in metastatic breast cancer. Gennari (2000): 6/9 responders *SUV levels as proportion of baseline* Average SUV drop after one cycle 18% vs no fall in remaining three	See above Heterogeneous treatment regimens, PET interpretation methods, response criteria and analytical methods Quality of studies generally poor.

	Population	Index test	Results	Comments
			Schwarz (2005): 6/11 responders *SUV levels as proportion of baseline* Responders: 72% ± 21% after cycle 1, 54% ± 16% after cycle 2 Non-responders: 94% ± 19% cycle 1, 76% ± 9% cycle 2 *Conclusion*: There are insufficient data to recommend PET scan for monitoring of treatment response. More studies are needed looking at response in lymph nodes.	
Systematic reviews				
Bourguet 2006	Patients with locally advanced breast cancer	**FDG-PET** Reference standard: histology and mammo/ US/ scan	**1 primary study (Kim 2004; n=50)**: see above. *Standard*: no standard applicable *Recommendation*: PET scans could predict response to neoadjuvant chemotherapy in locally advanced breast cancer.	Same study included in NCCHTA 2007

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Schwarz-Dose 2009	104 patients with newly diagnosed large (≥ 3 cm) or locally advanced non inflammatory breast cancer who participated in a prospective RCT comparing 2 regimens of preoperative chemotherapy	Sequential **FDG-PET** for monitoring response to neo-adjuvant chemotherapy in breast cancer Reference standard: histo-pathologic response after completion of chemotherapy	Histopathologic response after the first cycle of chemotherapy	According to the various tresholds for relative decrease in FDG Uptake (SUV) to predict histopathologic response after the first cycle of chemotherapy (from 20% to 50%); n=69: Se: from 67% to 93% Sp: from 22% to 70% After the first cycle, SUV decreased by 51% ± 18%in histopathologic responders, compared with 37% ± 21% in non responders (p<.01). An additional decrease of 63% ± 19% from baseline was observed after the second cycle in responders, versus 48% ± 19% in non responders (p<.01).	Prospective multicenter study 2 arms of trt The analysis grouped all patients Possibility of disease progression bias and review bias: unclear

	Population	Index test	Outcome	Results	Comments
				After a single cycle of chemotherapy, FDG-PET predicted pCR (specimens with no residual invasive tumor) with a sensitivity of 90% and specificity of 74%. Conclusion: results suggest a potential clinical application of FDG-PET in guiding systemic therapy of patients with locally advanced breast cancer on the basis of pre-therapy tumour metabolic activity. FDG-PET could be particularly helpful in identifying metabolically active tumours that may gain the most benefit from neo-adjuvant chemotherapy. This hypothesis needs additional evaluation in prospective trials.	
Berriolo A 2007	47 women with non-metastatic, non inflammatory, large or locally advanced breast cancer receiving different regimens of preoperative chemotherapy	**FDG-PET** for monitoring response to neo-adjuvant chemotherapy in breast cancer Reference : histopathologic response after completion of chemotherapy	Histopathologic response after the first cycle of chemotherapy	Early changes induced by neoadjuvant chemotherapy in tumour FDG uptake are highly predictive of the pathological response in breast cancer. FDG-PET allows adequate differentiation between pCR and non-pCR patients with both high sensitivity (91%) and high specificity (86%). Conclusion: after one course of neoadjuvant chemotherapy the reduction in FDG uptake is an early and powerful predictor of pCR. Recommendation: FDG PET uptake may improve patient management by avoiding ineffective chemotherapy or supporting the decision to continue dose-intensive preoperative chemotherapy in responding patients.	Prospective diagnostic study Possibility of disease progression bias and review bias: unclear

Prognosis

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Cachin F 2006	47 women with metastatic breast cancer were treated with a maximum of three cycles of high-dose chemotherapy	**FDG-PET**	Prediction of survival	Median follow-up : 87 months Median survival : 19 months The FDG-PET result was the most powerful and independent predictor of survival; patients with a negative post-treatment FDG-PET had a longer median survival than patients with a positive FDG-PET (24 months v 10 months; p < 0.001). The relative risk of death was higher in patients with FDG-PET-positive disease (RR: 5.3), prior anthracycline treatment (RR: 3.3), or with visceral metastasis (RR: 2.4). Conclusion: A single FDG-PET study performed after completion of high-dose chemotherapy for metastatic breast cancer can powerfully stratify for survival.	Prospective prognostic stratification study
Emmering J 2008	40 patients who were treated with neoadjuvant chemotherapy for locally advanced breast cancer (LABC)	**FDG-PET** Standard: post-operative histopathology	Prediction of overall survival (OS) and disease free survival (DFS)	Median follow-up: 60 months (range 15–94). Median time to progression was 26 months (range 14–90 months) Preoperative PET (HR 4.09; 95% CI 1.26–13.31; P = 0.02) was a better indicator for DFS than histopathological examination (HR 2.52; 95% CI 0.77–8.23; P = 0.13). In predicting OS, both PET (HR 2.77; 95% CI 0.66–11.66; P = 0.16) and histopathology (HR 6.53; 95% CI 0.80–	Prospective observational study Follow-up: check-ups once every 2 months in the first year after treatment, every 3 months in years 2–5 and every 6 months after year 5. Standard hematological and biochemical laboratory testing (including CA 15.3) was carried out every 6 months and a mammography of the contralateral breast every 12

	Population	Index test	Outcome	Results	Comments
				53.14; P = 0.08) were non significant predictors. Multivariate Cox regression revealed no added value of histopathology versus PET results. Conclusion: enhanced 18FDG uptake in the primary tumour of LABC patients after NC was associated inversely with DFS. The data suggest that to predict DFS, a preoperative 18FDG–PET is at least as good as postoperative histopathology and possibly better.	months. Computed tomography scan or US of the liver, bone scan or chest X-ray were carried out on indication only. Kaplan–Meier curves for DFS were plotted with PET and histopathology separately; a Cox proportional hazard regression analysis was carried out to control for possible confounders. Finally, a Cox regression analysis was carried out with both histopathology and PET as independent variables in one model. The same procedure was followed with OS.

ESOPHAGEAL CANCER

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with esophageal cancer	**FDG-PET** Objective: Diagnosis of primary tumour	One SR identified (MSAC 2001), already included in previous KCE report. Only patients with esophageal cancer included (no calculation of specificity possible). Reference standard: histopathology. No additional primary studies found. Conclusion: Diagnostic accuracy studies show accuracy, but unlikely to be used without biopsy in the UK	Good-quality HTA *Search date*: Aug 2005 *Databases*: Medline, EMBASE, Cochrane Library, HTA database, DARE, individual contacts through INAHTA No meta-analysis

Staging

	Population	Index test	Results	Comments
HTA reports				
MSAC 2008	Patients with esophageal cancer	**FDG-PET** **FDG-PET/CT**	Four HTA reports identified between 1999 and 2007 (NCCHTA 2007, AHTAPol 2006, KCE 2005, MSAC 2001). One additional SR identified (Westerterp 2006). Three primary studies identified (Meyers 2007, van Westreenen 2007, Stahl 2005) published since NCCHTA 2007. All suffered from partial verification (no 2x2 table possible).	Good-quality HTA *Search date*: Dec 2007 *Databases*: Medline, EMBASE, Pre-Medline, Cochrane Library Controlled Clinical Trials Registry, Current Contents No meta-analysis
NCCHTA 2007	Patients with esophageal cancer	**FDG-PET** <u>Objective</u>: Locoregional and distant staging	Two systematic reviews identified (BCBS HTA 2002, van Westreenen 2004), both included in previous KCE report. Population: biopsy-proven esophageal cancer. Reference standard: histopathology, clinical follow-up. Four additional primary studies found (Choi 2004, Heeren 2004, Liberale 2004, Sihvo 2004). Population: primary esophageal cancer. Reference standard: histopathology (or follow-up in 1 study). Three of these were diagnostic studies that confirmed the low sensitivity for locoregional lymph-node metastases (range 35-55%), but two showed slightly higher levels of sensitivity for distant metastases (71% and 88%), whereas one showed a lower level of 53% (or 64% when CT was added).	See above
Systematic reviews				
Van Vliet 2008	Patients with esophageal cancer	**FDG-PET** <u>Objective</u>: Preoperative staging	Reference standard: histopathology or follow-up. <u>Regional lymph node metastases</u>: Ten primary studies included, involving 424 patients (Flanagan 1997; Luketich 1997; Choi 2000; Lerut 2000; Wren 2002; Rasanen 2003; Yoon 2003; Heeren 2004; Sihvo 2004; Lowe 2005). Pooled sensitivity: 0.57 (95% CI 0.43–0.70) Pooled specificity: 0.85 (95% CI 0.76–0.95) Pooled log odds ratio: 1.71 (95% CI 1.22–2.20). It was not possible to assess whether publication bias was present, as the number of articles was too small (n=10). The differences between the SROC curves of EUS, CT, and FDG-PET for N staging were not statistically significant. The	Moderate-quality SR *Search date*: Jan 2006 *Databases*: Medline Search limited to English language *Meta-analysis* using random effects model. To estimate the relationship between sensitivities and specificities of each investigation, a random effects SROC analysis was performed. Software: STATA 8.0

	Population	Index test	Results	Comments
			relative diagnostic odds ratio of CT vs. EUS was 0.76 (95% CI 0.48–1.21; p=0.25) and of FDG-PET vs. EUS 0.95 (95% CI 0.54–1.67; p=0.86). Distant metastases: Nine primary studies included, involving 475 patients (Luketich 1997; Flamen 2000; Lerut 2000; Wren 2002; Rasanen 2003; Yoon 2003; Heeren 2004; Sihvo 2004; Lowe 2005). Pooled sensitivity: 0.71 (95% CI 0.62–0.79) Pooled specificity: 0.93 (95% CI 0.89–0.97) Pooled log odds ratio: 2.93 (95% CI 2.41–3.45) Assessment of publication bias was not possible. If the pooled sensitivities, specificities, and log odds ratios across tests were compared separately, we found higher values of FDG-PET for the detection of distant metastases compared to CT, although not statistically significant. Nevertheless, the SROC analysis showed that the diagnostic performance of FDG-PET was significantly higher than the diagnostic performance of CT (relative diagnostic odds ratio = 2.26 (95% CI 1.09–4.71), p<0.03), taking into account the inverse relationship between sensitivity and specificity and different test thresholds across the studies.	
Westerterp 2006	Patients with esophageal cancer	**FDG-PET** Objective: Preoperative staging	Four new primary studies identified since the SR of van Westreenen 2004 (Rasanen 2003, Kneist 2004, Sihvo 2004, Kato 2005b). Conclusions of authors: The 16 studies combined indicate that PET has hardly any additional value in the N-staging of esophageal cancer, while PET has additional value in M-staging of esophageal cancer in 3-20% of the patients.	Update of SR of van Westreenen 2004 (included in previous KCE report), however without specification about the used methodology and no meta-analysis with the new studies included.

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Schreurs 2008	Patients with esophageal cancer but without palpable cervical lymphadenopathy (n=180)	**FDG-PET** (n=125) Comparator: US of the neck Standard: Pathological outcome and/or clinical evidence of progressive cervical disease	N-staging (cervical region)	*Sensitivity*: PET 100% (7/7; 95%CI 59-100%), US 86% (6/7; 95%CI 42-100%) *Specificity*: PET 97% (115/118; 95%CI 93-99%), US 100% (118/118; 95%CI 97-100%) (NB: slightly discordant data in table 5 of article)	Retrospective study (the prospective study involving 55 patients did not have PET as index test). Consecutive patients. Differential verification. No definition of 'clinical evidence of progressive cervical disease'.
Little 2007	Patients with superficial adenocarcinoma of the esophagus (pTis [highgrade dysplasia] or pT1) undergoing esophagectomy (n=58)	**FDG-PET** (n=58) **FDG-PET/CT** (n=53; no results provided) Standard: Surgical findings, histopathology	N-staging M-staging	N-staging: *Sensitivity* 0% (0/6; 95%CI 0-46%) *Specificity* 94% (49/55; 95%CI 84-99%) M-staging: (no patients with distal metastasis) *Specificity* 95% (55/58; 95%CI 86-99%)	Unclear if prospective study.
Huguier 2006	Patients with cardio-esophageal cancer (n=28)	**FDG-PET** Standard: Histopathology, clinical follow-up	M-staging	*Sensitivity* 67% (4/6; 95%CI 22-96%) *Specificity* 100% (22/22; 95%CI 85-100%) PET modified the surgical strategy in 2 patients (7%).	Prospective study. Differential verification. No details on imaging tests used for follow-up. No 2x2 table possible for diagnosis.
Kato 2005a	Patients with thoracic esophageal cancer who underwent FDG-PET and bone scan within 1 month (n=44)	**FDG-PET** Comparator: Tc-99m bone scintigraphy Standard: Histology and/or clinical/imaging follow-up (including bone scintigraphy and FDG-PET)	M-staging (bone)	*Sensitivity*: PET 92% (12/13; 95%CI 64-100%), bone scan 77% (10/13; 95%CI 46-95%) *Specificity*: PET 94% (29/31; 95%CI 79-99%), bone scan 84% (26/31; 95%CI 66-95%)	Unclear if prospective study. Incorporation bias. Differential verification.

Monitoring of treatment response

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with esophageal cancer treated with neaodjuvant treatment and eligible for curative surgery	**FDG-PET** Objective: Assessment of response after neoadjuvant therapy	One systematic review identified (Westerterp 2005), already included in previous KCE report. Standard: histopathology. One additional primary study identified (Wieder 2004), also included in previous KCE report. Standard: histopathology. Conclusion: Studies show that FDG-PET may be superior to CT and comparable or superior to EUS, but small residual masses may be missed.	See above
NCCHTA 2007	Patients with esophageal cancer treated with neaodjuvant treatment	**FDG-PET/CT** Objective: Assessment of response after neoadjuvant therapy	One primary study identified (Cerfolio 2005, see above). Standard: histopathology. PET was found to be more sensitive than CT and EUS for assessment of treatment response after treatment (87% vs. 27% and 20%).	See above
Systematic reviews				
Westerterp 2006	Patients with esophageal cancer treated with neaodjuvant treatment	**FDG-PET** Objective: Assessment of response after neoadjuvant therapy	Two new primary studies identified since the SR of van Westreenen 2004 (Wieder 2004, Swisher 2004). Conclusions of authors: The diagnostic accuracy of FDG-PET in assessing response to treatment was similar to the accuracy of EUS, but significantly higher than that of CT.	See above

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Higuchi 2008	Patients with locally-advanced esophageal cancer who received neoadjuvant therapy (n=50)	**FDG-PET** (prior to neoadjuvant treatment and preoperatively) Standard: Histopathology	Treatment response Overall survival (Kaplan-Meier)	Treatment response (post-treatment SUVmax <2.5): *Sensitivity*: 90% (18/20) *Specificity*: 90% (27/30)	Prospective study. Only univariate analysis for survival analysis.

	Population	Index test	Outcome	Results	Comments
Wieder 2007	Patients with locally advanced adenocarcinoma of the distal esophagus scheduled to undergo neoadjuvant chemotherapy (n=24)	**FDG-PET** (prior to therapy, 2 weeks after initiation of therapy, and preoperatively) Standard: Surgical findings and histopathology	Treatment response (all patients with less than 10% viable residual tumour cells [regression score, grade 1] were classified as responding) Survival (Kaplan-Meier)	Treatment response: Using cut-off of -33% for SUV decrease between first and second PET: *sensitivity* 100%, *specificity* 63%. Using cut-off of -63% for SUV decrease between first and third PET: *sensitivity* 75%, *specificity* 87%.	Unclear if prospective study (probably yes). No multivariate analysis for survival analysis.
Song 2005	Patients with potentially resectable locally-advanced esophageal cancer undergoing neoadjuvant CRT (n=32)	**FDG-PET** (before and after CRT) Standard: Surgical findings and histopathology	Diagnosis of pathologic complete response	*Sensitivity*: 95% (20/21) *Specificity*: 27% (3/11) (NB: positive cases defined as negative preoperative PET)	Prospective study. Originally inclusion of 74 patients, but no post-treatment PET in 42 patients because of transient overloaded schedules in the nuclear medicine department and economic problems. Exclusion of 5 patients from analysis because of negative pre-treatment PET. No information on blinding.
FDG-PET/CT					
Roedl 2008	Patients with adenocarcinomas of the esophagus who had undergone PET–CT before and after neoadjuvant CRT (n=51)	**FDG-PET/CT** Contrast-enhanced (before and after CRT) Standard: Histopathology	Treatment response (histopathologic response was defined as <10% viable cells in the postsurgical tumour specimen) Disease-free and overall survival (Kaplan-Meier)	Decrease of PET/CT-volume was the single best predictor of histopathologic response. With a threshold of a 63% decrease of tumour volume, PET/CT-volume was able to predict histopathologic response with a sensitivity of 91% and a specificity of 90% (no 2x2 table). Metabolic response (decrease of total lesion glycolysis ≥78%): *sensitivity* 90% (19/21; 95%CI 70-99%), *specificity* 93% (28/30; 95%CI 78-99%).	Retrospective study.

	Population	Index test	Outcome	Results	Comments
				The decrease of the PET/CT volume was demonstrated to be the best predictor of disease-free survival and overall survival. Disease-free survival: metabolic responders 29.4 months (95%CI 26.3-32.6) vs. metabolic non-responders 16.0 months (95%CI 13.7-18.3; chi-square 25.5, p < 0.001; Log-Rank test). Overall survival: 34.1 months (95%CI 31.4 -36.8) vs. 21.8 months (95%CI 19.4-24.2) (chi-square 14.9, p < 0.001; Log-Rank test).	
Erasmus 2006	Patients with biopsy-proven esophageal cancer treated with neoadjuvant CRT and esophagectomy (n=56)	**FDG-PET/CT** Standard: Histopathology	Treatment response (complete response = no viable cancer in resection specimen)	**Discordant data in text!** Detection of complete response: Visual analysis: *Sensitivity* 54% (7/13; 95%CI 25-81%) *Specificity* 45% (13/29; 95%CI 26-64%) Semi-quantitative analysis (complete response = SUV <4): No 2x2 table possible.	Retrospective study. Consecutive patients. Forty-two patients included in analysis.

Detection of recurrence

Primary studies

	Population	Index test	Outcome	Results	Comments
Guo 2007	Patients with suspected recurrence (questionable symptoms or signs, equivocal diagnosis by CT, EUS, MRI, or barium swallow) after definitive treatment of esophageal squamous cell carcinoma (n=56)	**FDG-PET/CT** Standard: Histology or follow-up (serial imaging)	Detection of recurrent disease Overall survival (Kaplan-Meier)	Detection of recurrence: *Sensitivity* 96% (95%CI 85-99%) *Specificity* 55% (95%CI 23-83%) Overall survival: In multivariate survival analysis, therapeutic modality (HR 0.437; p=0.044), SUV (HR 1.071; p=0.029), and disease status on PET/CT (HR 2.430; p=0.045) were independent significant prognostic predictors for overall survival.	Unclear if prospective study. Differential verification. No information on what tests were included in follow-up.

Radiotherapy planning

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with esophageal cancer	**FDG-PET** Objective: Radiotherapy planning	Two primary studies identified (Konski 2005, Vrieze 2004), involving 55 patients in total. When PET was used in RT planning, it resulted in different target volumes.	See above

Prognosis

	Population	Index test	Results	Comments
HTA reports				
MSAC 2008	Patients with esophageal cancer	**FDG-PET** **FDG-PET/CT**	Five primary studies identified (Konski 2007, Chatterton 2006, Rizk 2006, van Westreenen 2005, Hong 2005) published since NCCHTA 2007.	See above.

Primary studies

	Population	Index test	Outcome	Results	Comments
Roedl 2009	Patients with esophageal squamous cell carcinoma, who received PET/CT scans both prior and after neoadjuvant chemoradiotherapy (n=49)	**FDG-PET/CT**	Disease-free survival (Kaplan-Meier)	The decrease of the metabolic tumour diameter between pre- and post-treatment PET/CT was the single best predictor of tumour-free survival. However, the accuracy of predicting survival was even higher when using the decrease of the "diameter-SUV index" as the metabolic criterion for treatment response. Metabolic responders (defined as a decrease of the diameter–SUV index by more than 55%) had a mean disease-free survival of 32 months, whereas metabolic non-responders (decrease of the diameter–SUV index by equal or less than 55%) had a mean disease-free survival of 16 months (Chi-square 32.3, $p<0.001$; Log Rank test).	Unclear if prospective study. No 2x2 table possible for preoperative evaluation of treatment response. Possible overlap with Roedl 2008!
Cheze-Le Rest 2008	Patients with newly diagnosed esophageal cancer (endoscopic biopsy) (n=52)	**FDG-PET**	Overall survival (Kaplan-Meier)	After multivariate analysis, SUVmax >9 (median survival 13 vs. 26 months, p=0.02) and FDG-positive lymph nodes (median survival 12 months if >1 positive node vs. 25 months if only 1 positive node) were found as independent predictors of poor outcome.	Prospective study. Consecutive patients. Median follow-up: 32 months (range 24-52 months).

	Population	Index test	Outcome	Results	Comments
Chung 2008	Patients with newly diagnosed esophageal cancer (n=100)	**FDG-PET** Comparator: 201Tl-SPECT	Disease-free survival	Survival analysis revealed tumour size, 201Tl negative tumors, FDG-negative tumours, delayed 201Tl T/M ratio, RI, stage, and FDG-pSUV to be significant univariate predictors for disease-free survival. Multivariate survival analysis showed only stage (p = 0.02) to be a significant independent prognostic predictor.	Unclear if prospective study. Consecutive patients. Mean follow-up: 20.3 ± 16.4 months (range 0-62.7 months).
Makino 2008	Patients with node-positive esophageal squamous cell carcinoma undergoing neoadjuvant chemotherapy (n=38)	**FDG-PET**	Disease-free survival	Multivariate analysis identified the number of PET-positive lymph nodes (p=0.018, HR=5.464) and PET response for primary tumour (p=.0015, HR=4.620) and for lymph nodes (p=0.028, HR=3.854) as independent prognostic predictors.	Retrospective study. Median follow-up after surgery: 31.2 months.
Omloo 2008	Patients with histologically proven cancer of the thoracic esophagus (type I) or cancer of the gastric cardia substantially involving the distal esophagus (type II) without evidence of distant metastases and/or locally irresectable disease (n=125)	**FDG-PET** Comparator: EUS	Disease-specific survival (Kaplan-Meier)	Patients with a high SUV (>0.27) had a significantly worse disease-specific survival compared with patients with a low SUV (OR 1.1; 95%CI 0.7-1.9; p=0.04). However, multivariate analysis of identified only EUS T-stage to be of independent prognostic significance.	Prospective study. No exact information on follow-up.
Roedl 2008	See above: Monitoring of treatment response.				
Kim 2007	Patients with resectable locally-advanced esophageal cancer undergoing preoperative CRT and surgery (n=62)	**FDG-PET(/CT)** Before and after CRT	Overall survival Disease-free survival	By multivariate analysis, complete metabolic response by FDG-PET was significantly associated with better DFS and OS (p=0.006 and p=0.033, respectively). The variables associated with pre-CRT PET scan were not predictive of survival.	Prospective study. Median follow-up: 19.3 months (range 3.9-57.1 months). Mix of FDG-PET and FDG-PET/CT (no figures provided).
Blackstock 2006	Patients with locally-advanced esophageal cancer undergoing neoadjuvant CRT (n=110)	**FDG-PET**	Overall survival	Multivariate analysis showed the following factors to be associated with survival: patients with non-adenocarcinoma histology (OR 33.4; 95%CI 2.25-493.9; p=0.01), patients without resection (OR 4.14; 95%CI 1.01-17.04; p=0.05), and patients with suggested distant metastasis after FDG-PET staging (OR 4.22; 95%CI 1.01-17.66; p=0.05).	Prospective study. Consecutive patients. Median follow-up: 13.8 months.

	Population	Index test	Outcome	Results	Comments
Cerfolio 2006	Patients with biopsy-proven apparently resectable esophageal cancer (n=89)	**FDG-PET**	Overall survival	Variables that were found to be independent predictors of survival by Cox hazards regression analysis were TNM staging (p=0.032) and maxSUV (p=0.014). Patients with a maxSUV >6.6 had a significantly worse survival (31% vs. 89%, p<0.001).	Retrospective analysis of prospective database. No detailed information on duration of follow-up.
Choi 2006	Patients with newly diagnosed esophageal squamous cell carcinoma who underwent esophagectomy (n=51)	**FDG-PET**	Disease-free survival Overall survival	Multivariate analysis revealed that the pathologic stage (HR 7.88; 95%CI 1.98-31.5; p=0.003) and number of PET-positive nodes (0, 1, 2, or ≥3) (HR 1.93; 95%CI 1.07-3.48; p=0.03) were independent significant prognostic predictors for overall survival.	Unclear if prospective study. Consecutive patients. Mean follow-up: 36 +/- 27 months.
Duong 2006	Patients with esophageal cancer undergoing neoadjuvant CRT (n=53)	**FDG-PET**	Overall survival (Kaplan-Meier)	Complete metabolic response (CMR) was strongly predictive of survival (p<0.008) on multivariate analysis. Patients with CMR on PET had statistically superior survival compared with those who had an incomplete metabolic response (78% vs. 33% 2-year survival rate). The relative risk of death for non-CMR patients was increased 5.75-fold compared with the CMR group (95%CI 1.94–17.05, p<0.001).	Prospective study. Consecutive patients. No 2x2 table possible for preoperative evaluation of treatment response.
Ott 2006	Patients with biopsy-proven adenocarcinoma of the distal esophagus or cardia with or without metastases in local lymph nodes, undergoing neoadjuvant treatment (n=65)	**FDG-PET** (before treatment and on day 14 of the first chemotherapy cycle) Standard: Histopathology	Treatment response (less than 10% viable tumour cells) Overall survival (Kaplan-Meier) Recurrence	Survival: 3-year survival rate: 35% for metabolic non-responders vs. 70% for responders (p=0.019). In a proportional hazards model the risk of death for patients with a metabolic response was 34% of the patients without a metabolic response (p=0.019). Recurrence: Multivariate analysis demonstrated that metabolic response was the only factor predicting recurrence (p=0.018) in the subgroup of completely resected (R0) patients.	Prospective study. Median follow-up: 42 months (range 26-67 months). Eight patients excluded from analysis. Partial verification for treatment response (results not presented).

GASTRIC CANCER

	Population	Index test	Results	Comments
HTA reports				
MSAC 2008	Patients with gastric cancer	**FDG-PET** **FDG-PET/CT**	No HTA reports identified. No primary studies reporting the accuracy of PET or PET/CT following laparoscopy (with or without CT) were identified. No primary studies reporting the therapeutic impact of PET in staging patients with gastric cancer were identified.	Good-quality HTA *Search date*: Dec 2007 *Databases*: Medline, EMBASE, Pre-Medline, Cochrane Library Controlled Clinical Trials Registry, Current Contents No meta-analysis

Primary studies

	Population	Index test	Outcome	Results	Comments
Ott 2008	Patients with locally-advanced gastric cancer receiving neoadjuvant treatment (n=71)	**FDG-PET** (before neoadjuvant treatment and preoperatively) Standard: Histopathology	Treatment response (histopathologic response: <10% residual tumour cells) Overall survival	Treatment response: (metabolic response = decrease of >35% of SUV) *Sensitivity*: 69% (11/16; 95%CI 41-89%) *Specificity*: 82% (27/33; 95%CI 65-93%) Overall survival: Multivariate analysis revealed metabolic response as the only significant presurgical predictor for survival (p=0.045; RR 0.39, 95%CI 0.16-0.98).	Prospective study. Median follow-up for surviving patients: 56.0 months (range 35-104 months). Only 49 patients assessable for metabolic response.
Sun 2008	Patients with gastric cancer previously treated with surgery (n=23)	**FDG-PET/CT** Standard: Histopathology or clinical follow-up	Detection of recurrent disease	*Sensitivity*: 86% (12/14; 95%CI 57-98%) *Specificity*: 78% (7/9; 95%CI 40-97%) Clinical treatment decisions were changed in 7 patients (30%) after introducing PET/CT into their conventional post-operative follow-up program.	Retrospective study. Mix of patients with suspicion of recurrence (imaging, history, clinical exam; n=12) and patients undergoing routine postoperative follow-up (n=11). Differential verification. Unclear if PET/CT was part of follow-up.

	Population	Index test	Outcome	Results	Comments
Di Fabio 2007	Patients with advanced gastric or GOJ cancer treated with chemotherapy plus cetuximab (n=22)	**FDG-PET** (at baseline and 6w after the start of treatment) Standard: Clinical follow-up and CT	Treatment response (complete clinical response: complete disappearance of target tumour lesions both clinically and at CT)	Using a cut-off of -35% of SUV: *Sensitivity*: 83% (95%CI 62-104) *Specificity*: 75% (95%CI 45-105)	Prospective study. Twenty evaluable patients.

THYROID CANCER

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with suspicious thyroid nodules	**FDG-PET**	One primary study (Kresnik 2003; 43 patients with suspicious thyroid nodules). Reference standard: histopathology. Sensitivity 100%, specificity 63% for the detection of malignant nodules.	Good-quality HTA *Search date*: Aug 2005 *Databases*: Medline, EMBASE, Cochrane Library, HTA database, DARE, individual contacts through INAHTA No meta-analysis

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Sebastianes 2007	Patients with thyroid nodules and indeterminate cytological results (n=42)	**FDG-PET** Standard: Histopathology	Diagnosis of malignancy Number of prevented unnecessary thyroidectomies	*Sensitivity*: 100% (95%CI 72-100%) *Specificity*: 39% (95%CI 22-58%) Twelve of 31 (39%) unnecessary thyroidectomies would be avoided using preoperative FDG-PET.	Prospective study. Not clear if histopathologic results were interpreted in a blind fashion.
de Geus-Oei 2006	Patients with a palpable solitary thyroid nodule and inconclusive FNAB findings (n=44)	**FDG-PET** Standard: Histopathology	Diagnosis of malignancy Number of prevented unnecessary thyroidectomies	*Sensitivity*: 100% (95%CI 54-100%) *Specificity*: 66% (95%CI 49-80%) Twenty-five of 38 (66%) unnecessary thyroidectomies would be avoided using preoperative FDG-PET.	Prospective study. Not clear if histopathologic results were interpreted in a blind fashion.

Restaging (after treatment)

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with thyroid cancer	**FDG-PET**	Four primary studies found (Gotthardt 2004, Hsu 2002, Iwata 2004, Shiga 2001), including in total 92 patients. Gotthardt et al. included patients with metastatic medullary thyroid cancer. The 3 other (Asian) studies used PET a few weeks after therapy to detect malignant lesions. Mixed reference standards (other imaging in 2 studies, histopathology or 1311 uptake in 1 study, histopathology or other imaging in 1 study). Gotthardt 2004 (n=26): no 2x2 table provided Iwata 2004 (n=19): per-lesion-analysis Hsu 2002 (n=15): one false-negative Shiga 2001 (n=32): per-lesion-analysis In one study CT detected more tumour sites than PET (Gotthardt 2004). Based on a small number of patients/lesions, PET appeared to have slightly better accuracy than whole-body scan (WBS) with various tracers in two studies, but not in the third (but this used a low dose of FDG).	See above

Detection of recurrence

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with previously treated thyroid cancer	**FDG-PET**	Two systematic reviews (Hooft 2001, AHRQ 2002 [HTA]) identified, both included in previous KCE report. Five additional primary studies (Chen 2003, Frilling 2001, Gabriel 2004, Groheux 2005, Yeo 2001) evaluated patients with *well-differentiated thyroid cancer* who had previously undergone treatment (thyroidectomy + ablation). In 3 studies recurrence was suspected based on elevated thyroglobulin levels and negative 1311 WBS. In 2 studies it was unclear on what the suspicion was based. Mixed reference standard (exclusively histopathology in only 1 study). Four out of five studies showed PET sensitivity of at least 80%, with specificity ranging from 25 to 83%. Two additional primary studies (De Groot 2001, Szakáll 2002) evaluated recurrent *medullary thyroid cancer* (66 patients in total, previously treated with surgery, and having elevated calcitonin or CEA levels). Reference	See above

	Population	Index test	Results	Comments
			standard: histopathology or clinical/imaging follow-up. Sensitivity of PET varied between the studies (41% vs. 95%), but PET identified more lesions than other methods. In one study, PET led to correct surgical intervention in 8 out of 26 patients. It was noted that PET failed to detect lesions that were smaller than 1 cm.	
NCCHTA 2007	Patients with thyroid cancer	**FDG-PET/CT**	Two primary studies identified (Nahas 2005, Ong 2005) of PET/CT in suspected recurrence of differentiated thyroid cancer in patients with elevated thyroglobulin and negative 1311 scintigraphy (50 patients in total). Reference standard: histopathology (both studies suffering from partial verification?). Both seem to be highly selected populations with few true negatives, so specificity figures are unreliable. In the larger prospective study, PET/CT had a sensitivity of 66%.	See above

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Freudenberg 2008	Patients with previously treated differentiated thyroid carcinoma (surgery + ablation) and increasing pathological Tg values, Tg antibody titres or both, but whose cervical US showed no pathological findings (n=21)	**FDG-PET** **124I-PET** **124I-PET/CT** Comparator: CT Standard: Clinical follow-up, with (n=12) or without (n=9) histopathology from post-study operative procedures	Detection of recurrence Staging of recurrent disease	Detection of recurrence: *Sensitivity*: FDG-PET 80% (16/20; 95%CI 56-94%), 1124-PET 50% (10/20; 95%CI 27-73%), CT 75% (15/20; 95%CI 51-91%) *Specificity*: 100% for all tests (only 1 TN: 95%CI 3-100%)	Unclear if prospective study. Consecutive patients. Differential verification. Possibly incorporation bias (unclear if PET was included in clinical follow-up). No 2x2 table possible for staging (for patient-based analysis). No patient-based analysis for 124I-PET/CT.
Freudenberg 2007	Patients with previously treated differentiated thyroid cancer (surgery + ablation) and with clinical or serological signs of recurrent disease (ultrasonography +/- elevated Tg +/- elevated Tg antibodies) (n=36)	**FDG-PET** **FDG-PET/CT** Comparator: CT Standard: Histology and clinical follow-up	Detection of recurrence Staging of recurrent disease	Detection of recurrence: *Sensitivity*: PET 91% (20/22; 95%CI 71-99%), PET/CT 95% (21/22 ; 95%CI 77-100%), CT 77% (17/22 ; 95%CI 55-92%) *Specificity*: PET 79% (11/14; 95%CI 49-95%),	Retrospective study. Consecutive patients. Differential verification. Incorporation bias. No 2x2 table possible for staging.

	Population	Index test	Outcome	Results	Comments
				PET/CT 100% (14/14 ; 95%CI 77-100%), CT 71% (10/14 ; 95%CI 42-92%) Compared with CT, FDG-PET changed therapy management in 9/36 patients (25%). Compared with CT+FDG-PET, FDG-PET/CT resulted in a treatment modification in 5/36 patients (14%).	
Alzahrani 2006	Patients with previously treated differentiated thyroid carcinoma (surgery +/- ablation) and elevated Tg with negative iodine scan (n=42), elevated Tg +/- abnormal US findings (n=8)	**FDG-PET** Standard: Fine-needle aspiration (FNA), histopathologic examination of subsequent surgical specimens, or persistent elevation of Tg levels for more than 1 year after the FDG-PET scan was obtained	Detection of recurrence	*Sensitivity*: 62% (95%CI 46-76%) *Specificity*: 88% (95%CI 47-100%)	Retrospective study. Differential verification.
Choi 2006	Patients with previously treated differentiated thyroid carcinoma (surgery + ablation) and elevated Tg or anti-Tg antibody or suspicious lesions on imaging, with negative iodine scan (n=108)	**FDG-PET** Standard: Histopathology (n=28), clinical/imaging follow-up (n=35), unknown in 45 patients (stated to be in remission)	Detection of recurrence	*Sensitivity*: 94% (95%CI 85-98%) *Specificity*: 78% (95%CI 63-89%)	Retrospective study. Differential verification. Reference standard unclear in 45 patients!
Iagaru 2006	Patients with previously treated papillary thyroid cancer and elevated Tg in the presence of a negative anatomic study (CT, MRI, US) and iodine scan (n=21)	**FDG-PET** (FDG-PET/CT: n=8) Standard: Histopathology (n=16) or follow-up (n=5)	Detection of recurrence	*Sensitivity*: 88% (95%CI 64-99%) *Specificity*: 75% (95%CI 19-99%)	Retrospective study. Consecutive patients. Differential verification. Unclear if PET was part of follow-up.
Pryma 2006	Patients with previously treated (surgery) Hürthle cell thyroid cancer and elevated Tg, abnormal conventional imaging	**FDG-PET** Standard: Clinical/imaging follow-up	Detection of recurrence Survival	Detection of recurrence: *Sensitivity*: 96% (95%CI 79-100%) *Specificity*: 95% (95%CI 75-100%)	Retrospective study. Differential verification. No multivariate analysis for prognosis.

	Population	Index test	Outcome	Results	Comments
	findings, or high-risk histopathologic findings (n=44)	+/- histopathology	(Kaplan-Meier)	Survival: Patients with SUVmax ≥10 had 5-y all-cause survival of 64% compared with 92% in those with SUVmax <0 (p<0.01).	
FDG-PET/CT					
Finkelstein 2008	Patients with previously treated differentiated thyroid cancer (surgery + ablation) and suspicion of recurrence (elevated Tg, elevated Tg antibodies, clinical suspicion) with negative iodine scan (n=65)	**FDG-PET/CT** Standard: Histopathology, serial imaging/exams, Tg level	Detection of recurrence	*Sensitivity*: 98% (95%CI 88-100%) *Specificity*: 81% (95%CI 58-95%)	Retrospective analysis of prospective database. Differential verification. Unclear if PET was part of serial imaging.
Freudenberg 2007	See above in this table.				
Shammas 2007	Patients with previously treated differentiated thyroid cancer (surgery + ablation) and suspicion of recurrence (elevated Tg, elevated Tg antibodies, clinical suspicion) with negative iodine scan (n=61)	**FDG-PET/CT** Standard: Follow-up imaging (neck ultrasound, MRI, CT, and post-radioiodine treatment scanning); Tg levels; histologic examination of surgical specimens	Detection of recurrence	*Sensitivity*: 68% (95%CI 51-82%) *Specificity*: 83% (95%CI 61-95%) In 27 of the 61 patients (44%), FDG-PET/CT resulted in subsequent treatment changes.	Retrospective study. Consecutive patients. Differential verification.

Monitoring of treatment response

	Population	Index test	Results	Comments
HTA reports				
NCCHTA 2007	Patients with advanced thyroid cancer	FDG-PET	One primary studies found (Boerner 2002; 21 patients with advanced thyroid cancer treated with isotretinoin; reference standard: clinical response): non-significant trend towards lower FDG uptake at 3 months in tumours with better long-term outcome.	See above

Prognosis

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Robbins 2006	Patients with follicular cell-derived thyroid carcinoma (n=400)	**FDG-PET**	Overall survival (Kaplan-Meier)	FDG-status (RR 7.69; 95%CI 2.17–24.4) and the number of FDG lesions (RR 1.1; 95%CI 1.08 – 1.15) significantly correlated with survival.	Retrospective prognostic study.

PANCREATIC CANCER

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with suspected pancreatic cancer	**FDG-PET**	Seven prospective studies identified including 479 participants, and with reference standard histology/biopsy or clinical follow-up (Casneuf 2007, Sperti 2007, Bang 2006, Maemura 2006, Ruf 2006, Lytras 2005, Lemke 2004). Outcome: diagnosis & staging (not separated). *Sensitivity* ranged from 73% to 97%, *specificity* ranged from 41% to 97%. Four prospective studies on primary diagnosis identified including 230 participants, and with mixed reference standards (Nishiyama 2005, van Kouwen 2005, Giorgi 2004, Rasmussen 2004). One retrospective study excluded from meta-analysis (Mansour 2006). *Sensitivity* ranged from 69% to 91%, *specificity* ranged from 65% to 100%.	Good-quality HTA *Search date*: 2003 - March 2008 *Databases*: Medline, EMBASE, Central, Scopus *Meta-analysis* using the DerSimonian and Laird random effects method. Software: RevMan software version 5.0.
AHRQ 2008	Patients with suspected pancreatic cancer	**FDG-PET/CT**	Three prospective studies identified including 193 participants (Casneuf 2007, Heinrich 2005, Lemke 2004). *Sensitivity* was 89% percent in all the individual studies, *specificity* ranged from 64% to 90.	See above.

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Singer 2007	Patients with mass-forming lesion of the pancreas (ultrasound, CT or MRI) (n=41)	**FDG-PET** Standard: Histology (biopsy, surgery, autopsy) (n=25) or clinical/imaging follow-up (n=16)	Diagnosis of primary tumour (n=41) (NB: including neuro-endocrine tumours and common bile duct tumours)	*Sensitivity*: 86% (95%CI 65-97%) *Specificity*: 79% (95%CI 54-94%)	Not stated if prospective study. Consecutive patients. Differential verification. No explanation on which imaging studies were used during follow-up.
Sperti 2005	Patients with suspected cystic tumour of the pancreas (n=33) or intraductal papillary mucinous tumour (n=17) (CT, CA 19-9, MRI)	**FDG-PET** Standard: Pathology, biopsy, follow-up	Diagnosis of malignant lesions (n=50)	*Sensitivity*: 94% (16/17; 95%CI 71-100%) *Specificity*: 94% (31/33; 95%CI 80-99)	Prospective study. Differential verification. Unclear what imaging studies were included in follow-up.
FDG-PET/CT					
Schick 2008	Patients with solid pancreatic lesions of unknown dignity with a diameter of ≥10 mm (n=46)	**FDG-PET/CT** Comparators: EUS (n=45), ERCP (n=39), US (n=38) Standard: Histopathology (n=43) or clinical follow-up (n=3)	Detection of malignancy (n=46)	*Sensitivity*: PET/CT 89% (95%CI 71-98%) (EUS 81%, ERCP 87%, US 86%) *Specificity*: PET/CT 74% (95%CI 49-91%) (EUS 84%, ERCP 88%, US 88%) No significant differences with comparators. PET/CT revealed cervical lymphonodal metastasis from occult bronchogenic carcinoma and a tubular colon adenoma with intermediate dysplasia on polypectomy, respectively.	Prospective study. Differential verification. Unclear what imaging tests were included in follow-up. Not all patients received comparator tests.

Staging

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with pancreatic cancer	**FDG-PET**	See above in table on diagnosis: results on staging not separately discussed in 7 prospective studies. Apart from these, 2 primary studies identified specifically on staging (Nishiyama 2005, Wakabayashi 2008). Nishiyama 2005 (n=42; prospective; M-staging): Se 81%, Sp 88% Wakabayashi 2008 (n=53; retrospective; N- and M-staging): no full 2x2 tables provided	See above

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Strobel 2008	Patients with biopsy-proven pancreatic cancer (n=50)	**FDG-PET** **FDG-PET/CT** (non-enhanced and enhanced) Standard: Intraoperative findings, histologic findings, or clinical/imaging follow-up	Resectability (absence of distant metastases, arterial infiltration, and infiltration of organs other than the duodenum or stomach)	*Sensitivity*: PET 100% (95%CI 85-100%), PET/CT 100% (85-100%), cePET/CT 96% (78-100%) *Specificity*: PET 44% (95%CI 25-65%), PET/CT 56% (35-75%), cePET/CT 82% (62-94%)	Retrospective study. Consecutive patients. Differential verification. No information on what imaging tests are included in follow-up.
Huguier 2006	Patients with suspected pancreatic cancer (based on other morphologic imaging test) (n=22)	**FDG-PET** Standard: Histopathology, clinical follow-up	M-staging (n=22)	*Sensitivity*: 70% (95%CI 35-93%) *Specificity*: 83% (95%CI 52-98%)	Prospective study. Differential verification. No details on imaging tests used for follow-up. No 2x2 table possible for diagnosis.
Nishiyama 2005a	Patients with pancreatic cancer (histopathology: n=41; clinical/imaging follow-up: n=14)	**FDG-PET** (with delayed additional imaging) Standard: Histology/cytology, follow-up (including FDG-PET)	Nodal staging Detection of liver metastasis	Nodal staging: *Sensitivity*: 70% (14/20; 95%CI 46-88%) *Specificity*: 97% (34/35; 95%CI 85-100%) Detection of liver metastasis: *Sensitivity*: 61% (11/18; 95%CI 36-83%) *Specificity*: 100% (37/37; 95%CI 90-100%)	Unclear if prospective study. Differential verification. Incorporation bias. Included in AHRQ 2008, but only for the results on primary diagnosis.
FDG-PET/CT					
Strobel 2008	See above in this table.				

Prognosis
Primary studies

	Population	Index test	Outcome	Results	Comments
Lyshchik 2005	Patients with pancreatic cancer (n=65)	**FDG-PET**	Overall survival	Multivariate analysis showed that only three factors had an independent association with longer patient survival: female gender (p<0.01), TNM stage I–III (p<0.05) and RI>10% (p<0.01).	Prospective study. Consecutive patients. No detailed information on duration of follow-up. SUV1 = SUV at 1h after FDG injection. SUV2 = SUV at 2h after FDG injection. RI = retention index = 100% x (SUV2-SUV1)/SUV1

PRIMARY LIVER CANCER

Staging
Primary studies

	Population	Index test	Outcome	Results	Comments
Ho 2007	Patients with primary hepatocellular carcinoma (n=121)	**Dual-tracer (11C-ACT and FDG) PET/CT** Standard: Histopathology; biochemical evidence of increasing a-FP and clinical follow-up; 2 or more serial PET/CT studies in a 3- to 7-mo period with unequivocal evidence of progression; additional or follow-up radiologic evidence of bone and lung metastases	M-staging	*Sensitivity*: FDG-PET/CT 79% (78/99) ACT-PET/CT 64% (63/99) Dual tracer 98% (97/99) *Specificity*: FDG-PET/CT 91% (20/22) ACT-PET/CT 95% (21/22) Dual tracer 86% (19/22)	Retrospective study. Differential verification. Incorporation bias.
Yoon 2007	Patients with hepatocellular carcinoma (n=87)	**FDG-PET** Standard: Imaging (chest CT; WBBS or MRI of the bone) and clinical follow-up	M-staging	Lung metastasis: *Sensitivity* 100% (12/12) *Specificity* 84% (63/75) Lymph node metastasis: *Sensitivity* 100% (19/19) *Specificity* 94% (64/68) Bone metastasis: *Sensitivity* 100% (11/11) *Specificity* 100% (76/76)	Unclear if prospective study. Consecutive patients. Differential verification.

Prognosis

Primary studies

	Population	Index test	Outcome	Results	Comments
Paudyal 2008	Patients with hepatocellular carcinoma (n=31)	**FDG-PET**	Overall survival	Multivariate analysis showed that a high SUV (>2) (HR 1.49, 95%CI 1.03-2.15; p=0.03) and lymph node metastasis (HR 0.05; 95%CI 0.0-0.06; p=0.04) were unfavourable prognostic factors.	Unclear if prospective study. No exact information on duration of follow-up
Seo 2007	Patients with hepatocellular carcinoma who underwent curative resection (n=70)	**FDG-PET**	Overall survival Disease-free survival	In multivariate analysis, a high a-FP level (RR 5.46, p=0.003; RR 8.78, p=0.006) and highTNR (tumour to non-tumour SUV ratio) (RR 1.3, p=0.03; RR 1.6, p=0.02) were independent predictors of postoperative recurrence and overall survival.	Prospective study. Mean follow-up: 596 days (range 75-1125 days).

CERVICAL CANCER

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with suspected cervical cancer	**FDG-PET**	One primary study identified on primary diagnosis and detection of recurrence (Chang 2005: n=219; FDG-PET). Reference standard: histology/biopsy or clinical follow-up. Lesion-based analysis.	Good-quality HTA *Search date*: 2003 - March 2008 *Databases*: Medline, EMBASE, Central, Scopus *Meta-analysis* using the DerSimonian and Laird random effects method. Software: RevMan software version 5.0.

Staging

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with cervical cancer	**FDG-PET**	Several primary studies identified: - 10 on initial staging (Chou 2006, Hope 2006, Lin 2003, Ma 2003, Park 2005, Roh 2005, Tran 2003, Unger 2005, Wright 2005, Yen 2003). Reference standard: histology/biopsy in 8 studies (2 studies also with clinical/imaging follow-up) - 1 on staging of primary and recurrent disease (Wong 2004) - 1 on initial staging and detection of recurrence (Grisaru 2004) Overall (7 studies, n=468): Se 10-100%, Sp 90-100%	See above

	Population	Index test	Results	Comments
			Chou 2006 (n=60; N-staging): Se 10%, Sp 94% Hope 2006 (n=58): incorrect 2x2 table provided Park 2005 (n=36; N-staging): Se 43%, Sp 100% Roh 2005 (n=59; N-staging): per-lesion analysis Wright 2005 (n=54; N-staging): pelvic LN Se 52%, Sp 90%, PALN Se 25%, Sp 97% Unger 2005 (n=14; N-staging): Se 29%, Sp 100% Wong 2004 (n=9; M-staging): Se 100%, Sp 100% Lin 2003 (n=14; N-staging): Se 86%, Sp 94% Ma 2003 (n=104; N-staging): Se 82%, Sp 97% Tran 2003 (n=186; N-staging): Se 100%, Sp 100% Yen 2003 (n=135): per-lesion analysis	
AHRQ 2008	Patients with cervical cancer	**FDG-PET/CT**	Several primary studies identified: - 5 on initial staging (Amit 2006, Choi 2006, Loft 2007, Sironi 2006, Yildirim 2008). Reference standard: histology/biopsy in 4 studies (2 studies also with clinical follow-up) - 1 on staging of primary and recurrent disease (Bjurberg 2007) Yildirim 2008 (n=16; N-staging): Se 50%, Sp 83% Bjurberg 2007 (n=17; M-staging): Se 94%, Sp NA Loft 2007 (n=119) : PALN Se 100%, Sp 99% ; M-staging Se 100%, Sp 94% Amit 2006 (n=75; extra-cervical lesions): Se 60%, Sp 94% Choi 2006 (n=22; N-staging): lesion-based analysis Sironi 2006 (n=47; N-staging): lesion-based analysis	See above
Systematic reviews				
Bourguet 2006	Patients with cervical cancer	**FDG-PET**	One systematic review (Havrilesky 2005) and 2 new primary studies (Park 2005, Yen 2003; see above) identified. No change in recommendation: FDG-PET can be proposed to ameliorate the nodal staging in cervical cancer. Havrilesky 2005: included 13 studies (1999-2004) PALN (4 prospective studies, n=136): Se 84% (68-94%), Sp 95% (89-98%) Pelvic LN (4 studies, n=162): Se 79% (65-90%), Sp 99% (96-99%)	Moderate-quality SR Update of a previous systematic review (2003), as a basis for the development of a CPG Literature search in Medline (2003-November 2005) + OVID alerts Language restrictions: French and English
Selman 2008	Patients with a primary presentation of cervical cancer of any histological type or stage	**FDG-PET** Outcome: Nodal staging	Eight primary studies identified on PET (Rose 1999a, Rose 1999b, Kuhnel 2001, Reinhardt 2001, Belhocine 2002, Yeh 2002, Lin 2003, Roh 2005). Pooled sensitivity: Sentinel node biopsy (SNB): 91.4 (95%CI 87.1–94.6)	Good-quality SR *Search date*: 2006 *Databases*: Medline, EMBASE, Cochrane Library, Medion

	Population	Index test	Results	Comments
		Reference standard: Histology	PET: 74.7 (63.3–84.0) MRI: 55.5 (49.2–61.7) CT: 57.5 (53.5–61.4) Pooled specificity: SNB: 100 (95%CI 99.6–100) PET: 97.6 (95.4–98.9) MRI: 93.2 (91.4–94.0) CT: 92.3 (91.1–93.5) No possibility to re-calculate Se and Sp of the individual studies based on the data of the meta-analysis.	No language restriction Random-effects *meta-analysis* of accuracy indices, meta-regression analysis to test the effect of study quality on diagnostic accuracy and to identify other sources of heterogeneity.

Detection of recurrence

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with suspected recurrent cervical cancer	**FDG-PET**	Thirteen primary studies identified (Chang TC 2004, Chang WC 2004, Chung 2006, Havrilesky 2003, Lin 2006, Ryu 2003, Sakurai 2006, Unger 2004, Van Der Veldt 2006, Yen 2006, Yen 2004, Chang 2005, Grisaru 2004). Reference standard: histology/biopsy or clinical follow-up in 10 studies, histology/biopsy in 3 studies. Chang TC 2004 (n=27): lesion-based analysis Chang WC 2004 (n=20): lesion-based analysis Chang 2005 (n=219): lesion-based analysis Chung 2006 (n=121): Se 96%, Sp 84% Havrilesky 2003 (n=28): lesion-based analysis Lin 2006 (n=26): peritoneum Se 57%, Sp 89%; bone Se 50%, Sp 96%; liver/spleen Se 100%, Sp 100%; lung Se 75%, Sp 100%; MLN Se 100%, Sp 88%; SLN Se 75%, Sp 95%; PALN Se 90%, Sp 94%; PLN Se 50%, Sp 100% Ryu 2003 (n=249): Se 90%, Sp 76% Sakurai 2006 (n=25) : lesion-based analysis Unger 2004 (n=44): per-PET analysis Van Der Veldt 2006 (n=38): no 2x2 table provided Yen 2006 (n=149): peritoneum Se 65%, Sp 98%; bone Se 100%, Sp 97%; liver/spleen Se 67%, Sp 99%; lung Se 92%, Sp 97%; MLN Se 100%, Sp 96%; SLN Se 81%, Sp 98%; PALN Se 88%, Sp 99%; PLN Se 83%, 98% Yen 2004 (n=55): peritoneum Se 88%, Sp 96%; bone Sp 98%; liver/spleen Se 100%, Sp 98%;	See above.

	Population	Index test	Results	Comments
			lung Se 78%, Sp 100%; MLN Se 100%, Sp 98%; SLN Se 85%, Sp 98%; PALN Se 88%, Sp 100%; PLN Se 91%, Sp 98% Conclusions of authors: FDG-PET is useful to detect or rule out recurrences, although there is some variation in the magnitude of the likelihood ratios across sites (3 prospective studies, n=230). The findings are consistent across each of the sites of recurrence in terms of being statistically significant, as well as for both prospective and retrospective (3 studies, n=396) designs.	
AHRQ 2008	Patients with suspected recurrent cervical cancer	**FDG-PET/CT**	Two primary studies identified (Chung 2007, Sironi 2007). Reference standard: histology/biopsy or clinical follow-up. Chung 2007 (n=52): Se 90%, Sp 81% Sironi 2007 (n=12) : Se 83%, Sp 100%	See above.
Systematic reviews				
Bourguet 2006	Patients with suspected recurrent cervical cancer	**FDG-PET**	One systematic review (Havrilesky 2005) and 3 new primary studies (Havrilesky 2003, Chang 2004, Unger 2004; see above) identified. No new recommendation. Havrilesky 2005: Three studies using FDG-PET in case of suspicion: Se 96% (87-99%), Sp 81% (58-94%) Two studies using systematic follow-up with FDG-PET: Se 92% (77-98), Sp 75% (69-80)	See above.

Primary studies

	Population	Index test	Outcome	Results	Comments
Kitajima 2009	Women who underwent treatment for histopathologically proven uterine cancer and with suspicion of recurrence (elevated levels of tumour markers, physical examination, abnormal CT and/or MR imaging findings, both elevated tumour marker levels and abnormal CT and/or MR imaging findings, or an abnormal Pap smear) (n=90)	**FDG-PET** **FDG-PET/CT** (contrast-enhanced) Comparator: CT Standard: Histopathology/ biopsy or clinical follow-up (including PET)	Detection of recurrent disease	*Sensitivity*: PET 80% (68–91%) PET/CT 91% (82–99%) CT 68% (54–82%) *Specificity*: PET 74% (61–87%) PET/CT 94% (86–100%) CT 87% (77–97%) The findings of PET/CT resulted in a change of management in 38 of the 90 patients (42%) which included initiating an unplanned treatment strategy (n=24), changing the treatment plan (n=8), and obviating the need	Prospective study. Consecutive patients. Differential verification. Incorporation bias. Uterine cervical cancer: n=50; endometrial cancer: n=40.

	Population	Index test	Outcome	Results	Comments
				for planned treatment (n=6).	
Kitajima 2008	Patients with previously treated cervical cancer and suspicion of recurrence (elevated levels of tumour markers and/or abnormal CT and/or MR imaging findings, physical examination, abnormal Pap smear) (n=52)	**FDG-PET** **FDG-PET/CT** (non-enhanced) Standard: Histopathology or clinical follow-up (including PET/CT in some patients)	Detection of recurrent disease	*Sensitivity*: PET 80% (95%CI 64-95%) PET/CT 92% (95%CI 81-100%) *Specificity*: PET 78% (62-94%) PET/CT 93% (83-100%)	Retrospective study. Consecutive patients. Differential verification. Incorporation bias.
van der Veldt 2008	Patients with histologically confirmed cervical cancer and primary treatment with curative intent, and suspicion of recurrence (clinically or radiologically) (n=40)	**FDG-PET** Standard: Histopathology or clinical follow-up	Detection of recurrent disease Staging of recurrent disease	Detection of recurrent disease: *Sensitivity*: 92% (95%CI 81-96%) *Specificity*: 93% (95%CI 71-100%) Staging of recurrent disease: Local recurrence: Se 100% (66-100%), Sp 97% (83-100%) Regional recurrence: Se 87% (60-98%), Sp 100% (86-100%) Distant metastasis : Se 75% (35-97%), Sp 100% (89-100%) Two experts reported that FDG-PET led to a better diagnosis and a beneficial change in management in 60% and 65% of cases, respectively.	Retrospective study. Consecutive patients. Differential verification.

Staging of recurrence

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with recurrent cervical cancer	**FDG-PET** **FDG-PET/CT**	Three primary studies identified (Bjurberg 2007, Wong 2004, Lai 2004). Reference standard: histology/biopsy or clinical follow-up. No conclusions provided. Bjurberg 2007 (n=15; M-staging): Se 92%, Sp 100% Lai 2004 (n=40): per-lesion analysis Wong 2004 (n=41): per-lesion analysis	See above.
Systematic reviews				
Bourguet	Patients with	**FDG-PET**	Two new primary studies (Lai 2004, Yen 2004; see above) identified.	See above.

	Population	Index test	Results	Comments
2006	recurrent cervical cancer		FDG-PET can be useful to decide on the therapeutic strategy.	

Primary studies

	Population	Index test	Outcome	Results	Comments
van der Veldt 2008	See above: Detection of recurrence.				

Prognosis

	Population	Index test	Results	Comments
Systematic reviews				
Bourguet 2006	Patients with cervical cancer	**FDG-PET**	One new primary study identified (Grigsby 2004). Normal FDG-uptake was associated with better 5-year survival. No change in recommendation.	See above.

Primary studies

	Population	Index test	Outcome	Results	Comments
Kidd 2008	Patients with stage Ib to IVa cervical cancer (n=72)	**FDG-PET/CT**	Risk of recurrence Overall survival	Cox proportional hazards modelling for pelvic recurrence showed that tumour volume, as determined by FDG-PET, was the most significant predictive factor of pelvic recurrence (p=0.0003) with a HR of 1.015 (95%CI 0.999-1.033). Tumour heterogeneity was the next most significant predictive factor of pelvic recurrence (p=0.0035) with a HR of 1.074 (95%CI 0.476-2.422). SUVmax did not remain a significant predictor (p=0.5713) in this model.	Prospective study. Mean follow-up for event-free patients: 21.1 months (range 7-61 months).
Yen 2008	Patients with cervical cancer and pelvic or para-aortic LN (PLN or PALN) metastasis detected by CT/MRI (n=70)	**FDG-PET**	Recurrence-free survival (RFS) Overall survival (OS)	SUVmax for PALN (dichotomized by 3.3) was significantly associated with OS (p=0.012) and marginally with RFS (p=0.078). The presence of SUVmax≥3.3 at PALN was significantly associated with both recurrence (5-year RFS; HR=4.52, 95%CI 1.73–11.80] and death (5-year OS; HR 6.04, 95%CI 1.97–18.57).	Prospective study. Consecutive patients. Mean follow-up for event-free patients: 47.6±13.4 months (range 28–72 months).
Schwarz 2007	Patients treated for advanced cervical cancer (concurrent	**FDG-PET/CT** (pre-treatment and 3 months	Progression-free survival	Progression-free survival: Progressive disease on 3-mo post-treatment PET: HR 32.57 (p<0.001; 95%CI 10.22-103.82)	Prospective study. Mean follow-up: 25 months (range 6-49 months).

	Population	Index test	Outcome	Results	Comments
	CRT) (n=92)	post-treatment)	Cause-specific survival	Partial metabolic response: HR 6.30 (95%CI 2.73-4.56; p<0.001) Pre-treatment lymph node status: HR 3.54 (p=0.03; 95%CI 1.54-8.09)	
Xue 2006	Patients with cervical cancer stage Ib1-IVb undergoing definitive radiotherapy +/- chemotherapy (n=96)	**FDG-PET** (pre-treatment and 3 months post-treatment)	Disease-free survival Overall survival	Lymph node metastasis on FDG-PET was found to be predictive of disease-free survival (p<0.0001). Both the SUV for FDG and FIGO Stage I disease were found to be marginally predictive of disease-free survival (p=0.055 and p=0.058, respectively).	Prospective study. Consecutive patients. Mean follow-up for event-free patients: 42.5 months (range 4-70 months). Blinded evaluation.

OVARIAN CANCER

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with (suspected) ovarian cancer	**FDG-PET** **FDG-PET/CT**	Three primary studies identified: one on primary diagnosis and staging (Castellucci 2007: FDG-PET/CT) and two on primary diagnosis alone (Kawahara 2004 [FDG-PET], Risum 2007 [FDG-PET/CT]). Reference standard: histology/biopsy. Castellucci 2007 (n=50): Se 87%, Sp 100% Risum 2007 (n=97): Se 100%, Sp 92% Kawahara 2004 (n=38): Se 78%, Sp 86%	Good-quality HTA *Search date*: 2003 - March 2008 *Databases*: Medline, EMBASE, Central, Scopus *Meta-analysis* using the DerSimonian and Laird random effects method. Software: RevMan software version 5.0.
AHRQ 2006	Women with adnexal mass	**FDG-PET** (amongst other imaging modalities: US, CT, MRI) Outcome: diagnosis of malignancy	Three primary studies identified on FDG-PET (Fenchel 2002, Grab 2000, Kawahara 2004). Reference standard: histopathology. Pooled sensitivity: 67% (95%CI 52-79%) Pooled specificity: 79% (95%CI 70-85) Author's conclusions: There is no evidence to support the superiority of any single modality, although FDG-PET appears inferior to the rest.	Good-quality HTA report *Search date*: September 2004 *Databases*: Medline, Cochrane Database of Systematic Reviews, references of review articles and meta-analyses *Meta-analysis* using SROC analysis and independently combined sensitivity and specificity values (software: Meta-Stat 0.6) HTA report with broader scope (management of adnexal mass): does not provide sufficient detail to allow confident estimation of the results

	Population	Index test	Results	Comments
Systematic reviews				
Bourguet 2006	Patients with suspected ovarian cancer	**FDG-PET**	One new primary study identified (Kawahara 2004, see above AHRQ 2008). No update of recommendations: no indication.	Moderate-quality SR Update of a previous systematic review (2003), as a basis for the development of a CPG Literature search in Medline (2003-November 2005) + OVID alerts Language restrictions: French and English

Primary studies

	Population	Index test	Outcome	Results	Comments
Yamamoto 2008	Women who were suspected to have ovarian cancer evidenced by US, MRI, and rising serum tumour markers (n=30)	**FDG-PET/CT** Standard: Histopathology	Diagnosis of malignancy (not exclusively ovarian cancer)	*Sensitivity*: 71% (10/14; 95%CI 42-92%) (including all patients with ovarian cancer) *Specificity*: 81% (13/16; 95%CI 54-96%)	Prospective study. No information on blinding.

Staging

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with ovarian cancer	**FDG-PET** **FDG-PET/CT**	Four primary studies identified: one on primary diagnosis and staging (Castellucci 2007: FDG-PET/CT), two on primary staging alone (Drieskens 2003 [FDG-PET/CT], Yoshida 2004 [FDG-PET]), and one on staging and detection of recurrence (Grisaru 2004: FDG-PET). Reference standard: histology/biopsy (or clinical follow-up in Grisaru 2004). Castellucci 2007 (n=50): no 2x2 table for staging Drieksens 2003 (n=13): lesion-based analysis Yoshida 2004 (n=15): lesion-based analysis Grisaru 2004 (n=18): no data on staging	See above
Systematic reviews				
Bourguet 2006	Patients with ovarian cancer	**FDG-PET**	Three new primary studies identified. No change in recommendation: no indication. FDG-PET: Sironi 2004 (n=31): lesion-based analysis Drieksens 2003 (n=13): lesion-based analysis	See above.

	Population	Index test	Results	Comments
			FDG-PET/CT: Yoshida 2004 (n=15): lesion-based analysis Picchio 2003 (n=25): Se 83%, Sp 92% (no 2x2)	

Detection of recurrence

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with suspected recurrent ovarian cancer	**FDG-PET** **FDG-PET/CT**	Twelve primary studies identified. Reference standard: histology/biopsy in Bristow 2003 and Bristow 2005, histology/biopsy or clinical follow-up in ten other studies. FDG-PET: 4 studies, n=223, Se 85-92%, Sp 78-100% Garcia-Velloso 2007 (n=86): Se 86%, Sp 78% Murakami 2006 (n=90): Se 91%, Sp 100% Takekamu 2005 (n=29): Se 85%, Sp 100% Grisaru 2004 (n=18): Se 92%, Sp 100% FDG-PET/CT: 9 studies, n=317, Se 73-100%, Sp 40-100% Sebastian 2008 (n=53): Se 97%, Sp 80% Chung 2007 (n=77): Se 93%, Sp 97% Kim 2007 (n=36): Se 73%, Sp 93% Thrall 2007 (n=39): Se 95%, Sp 100% Bristow 2005 (n=14): Se 77%, Sp 100% Hauth 2005 (n=19): Se 100%, Sp 100% Nanni 2005 (n=41): Se 88%, Sp 71% Pannu 2004 (n=16): Se 73%, Sp 40% Bristow 2003 (n=22): Se 83%, Sp 75%	See above.
Systematic reviews				
Bourguet 2006	Patients with suspected recurrent ovarian cancer	**FDG-PET**	One meta-analysis (Havrilesky 2005) and 3 new primary studies (Kim 2004, Nanni 2005, Takekuma 2005) identified. No change in recommendation: PET can be proposed in case of suspicion of recurrent disease taking into account that microscopic peritoneal disease can be the source of false negative results. Havrilesky 2005: included 10 primary studies (1993-2002) Five studies using systematic follow-up with FDG-PET: Pooled Se 54% (95%CI 39-69%), pooled Sp 73% (56-87%) Five studies using FDG-PET in case of clinical suspicion : Pooled Se 90% (82-95%), pooled Sp 86% (67-96%) Primary studies:	See above.

	Population	Index test	Results	Comments
			Nanni 2005 (n=41): Se 88%, Sp 71% Takekamu 2005 (n=29): Se 85%, Sp 100% Kim 2004 (n=55) : Se 82%, Sp 88%	

Primary studies

	Population	Index test	Outcome	Results	Comments
Kitajima 2008	Patients with previously treated ovarian cancer (primary cytoreductive surgery + chemotherapy) and suspected recurrence (elevated levels of CA-125, both elevated CA-125 levels and abnormal CT and/or MR imaging findings, abnormal CT and/or MR imaging findings, physical examination, and an abnormal Pap smear (n=132)	**FDG-PET/CT** (non-enhanced and enhanced) Comparator: Contrast-enhanced CT Standard: Histopathology/ biopsy or clinical follow-up of at least 6 months (including PET/CT in 30 cases)	Detection of recurrent disease	*Sensitivity*: nePET/CT 74% (95%CI 63.7-84.8%) cePET/CT 79% (95%CI 82.0-99.8%) ceCT 61% (95%CI 50.7-70.5%) *Specificity*: nePET/CT 91% (95%CI 84.0-97.8%) cePET/CT 90% (95%CI 81.7-97.5%) ceCT 85% (95%CI 76.1-93.5%)	Prospective study. Consecutive patients. Differential verification. Incorporation bias bias.

Staging of recurrence

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with recurrent ovarian cancer	**FDG-PET** **FDG-PET/CT**	Two primary studies identified (Picchio 2003 [FDG-PET/CT], Sironi 2004 [FDG-PET]). Reference standard: histology/biopsy. Picchio 2003 (n=25): lesion-based analysis Sironi 2004 (n=31): lesion-based analysis	See above.

Evaluation of treatment response

	Population	Index test	Results	Comments
Systematic reviews				
Bourguet 2006	Patients with ovarian cancer undergoing neoadjuvant treatment	**FDG-PET**	One new primary study identified (Avril 2005:FDG-PET) identified. No update of recommendations: no indication.	See above.

ENDOMETRIAL CANCER

	Population	Index test	Results	Comments
Systematic reviews				
Bourguet 2006	Patients with uterine cancer	**FDG-PET**	Two new primary studies found (Chao 2005, Saga 2003). No sufficient evidence to support recommendations. Chao 2005 (n=49): per-lesion-analysis Saga 2003 (n=21; detection of recurrence & evaluation of treatment response): Se 100%, Sp 91%	Moderate-quality SR, as a basis for a CPG (update of earlier report, that was included in the previous KCE report) *Search date*: November 2005

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Kitajima 2009	Women who underwent treatment for histopathologically proven uterine cancer and with suspicion of recurrence (elevated levels of tumour markers, physical examination, abnormal CT and/or MR imaging findings, both elevated tumour marker levels and abnormal CT and/or MR imaging findings, or an abnormal Pap smear) (n=90)	**FDG-PET** **FDG-PET/CT** (contrast-enhanced) Comparator: CT Standard: Histopathology/ biopsy or clinical follow-up (including PET)	Detection of recurrent disease	*Sensitivity*: PET 80% (68–91%) PET/CT 91% (82–99%) CT 68% (54–82%) *Specificity*: PET 74% (61–87%) PET/CT 94% (86–100%) CT 87% (77–97%)	Prospective study. Consecutive patients. Differential verification. Incorporation bias. Uterine cervical cancer: n=50; endometrial cancer: n=40.
Torizuka 2006	Patients with clinical stage I uterine corpus cancer, who underwent FDG-PET prior to surgery (n=22)	**FDG-PET** Comparator: MRI Standard: Histopathology	Assessment of myometrial infiltration	*Sensitivity*: PET 83% (5/6; 95%CI 36-100%) MRI 100% (6/6; 95%CI 54-100%) *Specificity*: PET 88% (14/16; 95%CI 62-98%) MRI 69% (11/16; 95%CI 41-89%)	Retrospective study. Possible selection bias, based on reference standard.
FDG-PET/CT					
Kitajima 2009	See above in this table.				
Chung 2008	Patients with previously	**FDG-PET/CT**	Detection of	*Sensitivity* 100% (12/12; 95%CI 74-100%)	Retrospective study.

	Population	Index test	Outcome	Results	Comments
	treated endometrial cancer (primary cytoreductive surgery followed by adjuvant treatment if necessary) and with suspected recurrence (symptoms, imaging, tumour markers, abnormal results on physical or cytological examination on routine surveillance, patient request of a surveillance PET/CT scan) (n=31)	Standard: Histology or follow-up (including PET/CT)	recurrent disease	*Specificity* 95% (18/19; 95%CI 74-100%)	Consecutive patients. Differential verification. Incorporation bias.
Kitajima 2008	Patients with histopathologically proven endometrial cancer scheduled for surgery (n=40)	**FDG-PET/CT** (non-enhanced) Standard: Histopathology	N-staging	*Sensitivity*: 50% (5/10; 95%CI 19-81%) *Specificity*: 87% (26/30; 95%CI 69-96%)	Prospective study. Consecutive patients.

RENAL CANCER

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with suspected RCC (diagnosis) or proven RCC (staging)	**FDG-PET** Reference standard: histology/biopsy or clinical follow-up	5 studies with small sizes for primary diagnosis only (Ak 2005) or for combined diagnosis and staging (Aide 2003, Chang 2003, Kang 2004, Kumar 2005) ***Prospective studies*** Diagnosis vs histology Ak (2005 (n=19): Suspected primary renal tumors based on conventional imaging techniques Se= 87% (95% CI 58%-98%) Sp= 75% (95% CI 22%-99%) Diagnosis and staging vs Histology/biopsy, follow-up (clinical	Studies included have small sample size Reference standards are different between studies without random assignment of patients A meta-analysis was conducted on the 3 retrospective studies (Chang 2003, Kang 2004, Kumar 2005); pooled negative and positive LR were provided but no results for pooled Se

	Population	Index test	Results	Comments
			course) (3-6 mo) Aide 2003 (n=35): Suspected RCC or RCC after radical or partial nephrectomy Se= 47% (95% CI 29%-65%) Sp= 80% (95% CI 30%-99%) ***Retrospective studies*** Diagnosis and staging vs Histology/biopsy, follow-up (clinical course) (3-6 mo) Chang 2003 (n=15): Histologically proven RCC and a solitary pulmonary lesion suspicious of lung metastasis Se 90% (55%-100%); Sp 80% (28%-99%) Kang 2004 (n=66): One year of follow-up or death due to rapidly progressive renal cell carcinoma within 1 year of the PET Se 60% (32%-84%); Sp 100% (16%-100%) Kumar 2005 (n=24): Suspected or known malignancies Se 89% (52%-100%); Sp 100% (3%-100%)	and Sp
Systematic reviews				
Bourguet 2006	Patients with suspected or known renal cell carcinoma	**FDG-PET** Reference standard: histology/biopsy or clinical follow-up	2 primary studies (Kang 2004; Kumar 2005); see AHRQ 2008.	Literature search: August 2006 Medline+ OVID Alerts+ EBM Websites

Staging

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with proven RCC (staging)	**FDG-PET** Reference standard: histology/biopsy or clinical follow-up (24	**Dilhuydy 2007** (n=24) : Patients with histologically proven renal cell carcinoma with metastatic disease awaiting a therapeutic decision for surgery, radiofrequency ablation, general specific treatment (immunotherapy) before surgery, or monitoring FDG-PET scans were performed after standard staging	Prospective study conducted in France on a small sample size (retrospective analysis). Interpretation of the reference standard was not blinded (review bias).

	Population	Index test	Results	Comments
		months)	(cerebral-thoracic-abdominal and pelvic CT scans and bone isotopes) for standard decision-making in metastatic RCC Sensitivity= 75% (95% CI : 47% - 92%) Specificity= 66% (95% CI : 12% - 98%) Treatment decision impact of FDG-PET imaging: overall, there were five changes (21%) to the management strategy. **In Kang study** (2004), 17 patients with suspicion of primary RCC who had not undergone nephrectomy performed a PET (17 scans) for staging and treatment orientation. 2 patients were accurately identified as having benign cysts by FDG-PET (Sp 100%); however, 6/15 (40%) disease positive individuals were not captured by FDG-PET imaging, yielding to a lower sensitivity (Se 60%) than abdominal CT demonstrated (Se 91.7% and Sp 100%). In Kumar (2005), 14 patients with metastatic renal tumors: no changes in treatment management were reported due to PET imaging. Sensitivity= 83% (95% CI: 58% - 96%) Specificity= Not defined (0/0)	There was more than one reference test (incorporation bias). + comparison between two index tests (not blinded): PET and CT Retrospective study which included patients for diagnosis, staging and restaging Retrospective study which included patients for diagnosis and staging Only patients with known renal masses were included PET imaging also formed a part of the reference standard in some instances (incorporation bias)
Systematic reviews				
Bourguet 2006	Patients with renal cell carcinoma	**FDG-PET**	No change since 2003: The place of PET in the initial staging of disease extension (evidence level C) remains to be determined in prospective studies.	Literature search: August 2006 Medline+ OVID Alerts+ EBM Websites Translation from French

Restaging

	Population	Index test	Results	Comments
HTA				
AHRQ 2008	Patients with RCC who had undergone nephrectomy (n=54)	**FDG-PET** Comparators: CT + bone scan Reference standard: Histology/biopsy, follow-up (clinical course) (12 mo)	In Kang study (2004), 54 patients undergone a PET (73 scans) for restaging. FDG-PET detected 64% of all soft tissue metastasis and 79% of bone metastasis. According to the localisation of metastases (lymph nodes, lung, liver, bone), FDG PET demonstrated: Se: 50-75% vs 77-100% (conventional imaging) Sp: 97-100% vs 73-98% (conventional imaging)	This study gave all diagnostic performance data (Se and SP) but no raw data (TP, FP, TN, FN) for detection of metastases.

Detection of recurrence

	Population	Index test	Results	Comments
Systematic reviews				
Bourguet 2006	Patients with known renal cell carcinoma	**FDG-PET** Reference standards: Histology/biopsy, follow-up (clinical course) (12 mo)	4 primary studies (Chang 2003, Kang 2004, Jadvar 2003, Majhail 2003) Jadvar 2003 (n=25): Non-diabetic patients with known or suspected metastatic RCC Se: 71% (95%CI: 48% - 88%) Sp: 75% (95% CI: 22% - 99%) Majhail 2003 (n=24): Histologically proven RCC undergoing surgical evaluation for possible resection of recurrent disease Se: 63% (95% CI: 45% - 79%) Sp: 100% (95%CI: 31%-100%) *Standard*: no standard applicable. *Option*: PET may be indicated in the search of local recurrences or distant metastasis (high positive predictive value) in case of suspected signs (pain, equivocal results of morphological imaging). However, a negative PET scan does not confirm the absence of recurrence (low negative predictive value) (evidence level B2).	Literature search: August 2006 Medline+ OVID Alerts+ EBM Websites Translation from French All studies are retrospective Reference standard is different for some patients (non-randomly assigned)

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET/CT					
Park 2009	63 patients with RCC were followed after surgical treatment (radical nephrectomy/ partial nephrectomy in 1 patient)	**FDG-PET/CT** Conventional imaging (plain chest radiography, abdominopelvic CT, whole body bone scan) Reference standard: histopathology or clinical follow-up	Detection of recurrence or metastases during follow-up	During a mean (range) of 24.3 (4–88) months of follow-up, 32 patients (51%) developed a local recurrence or distant metastases of RCC. FDG-PET/CT: Se: 94% (95%CI : 78% - 99%) Sp: 84% (95%CI : 65% - 94%) Conventional imaging: Se: 97% (95%CI : 82% - 99%) Sp: 81% (95%CI : 62% - 92%)	Retrospective study Diagnostic performance of PET/CT and conventional imaging (Se and Sp) are different than those reported in the paper

Monitoring of treatment response

	Population	Index test	Results	Comments
Systematic reviews				
Bourguet 2006	Patients with a renal cancer	**FDG-PET**	No evidence found. The utility of PET in the evaluation of therapy response requires assessment in prospective studies (experts' agreement).	Literature search: August 2006 Medline+ OVID Alerts+ EBM Websites Translation from French

TESTICULAR CANCER

Diagnosis

	Population	Index test	Results	Comments
Systematic reviews				
Bourguet 2007	Patients with testicular cancer	**FDG-PET**	No new primary studies found since the previous FNCLCC report (FNCLCC 2003).	Moderate-quality SR, as a basis for a CPG (update of earlier report, that was included in the previous KCE report) *Search date*: August 2006

Staging

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with testicular cancer	**FDG-PET**	One primary study identified (Lassen 2003). Forty-six patients included having undergone orchidectomy and negative postoperative conventional staging (abdominopelvic CT, chest X-ray, □FP and □HCG). Sensitivity 70%, specificity 100%, PPV 100%, NPV 92% and diagnostic accuracy 93%. Better diagnostic performance than conventional imaging, although not statistically significant (p<0.06).	Good-quality HTA *Search date*: 2003 - March 2008 *Databases*: Medline, EMBASE, Central, Scopus *Meta-analysis* using the DerSimonian and Laird random effects method. Software: RevMan software version 5.0.
Systematic reviews				
Bourguet 2007	Patients with testicular cancer	**FDG-PET**	One new primary study (Lassen 2003) found since the previous FNCLCC report: see AHRQ 2008.	See above

Primary studies

	Population	Index test	Outcome	Results	Comments
de Wit 2008	Patients with non-seminomatous germ cell tumours at an early stage (I and II) undergoing primary retroperitoneal lymph node dissection (n=72)	**FDG-PET** Comparator: CT Reference standard: histology (RPLND)	Nodal staging	FDG-PET vs. CT: Sensitivity 66% (95%CI 47-81%) vs. 41% (95%CI 24-59%) (p=0.038) Specificity 97% (95%CI 87-100%) vs. 95% (95%CI 83-99%) (NS) PPV 95% vs. 87% (NS) NPV 78% vs. 67% (p=0.05)	Prospective study Of the 87 enrolled patients, 15 were excluded: 14 because of observation without RPLND Consecutive patients?

Detection of recurrence

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with testicular cancer	**FDG-PET**	Two primary studies identified. Hinz 2008 (n=20): Se 100% (95%CI 29%–100%), Sp 47% (95%CI 23–72%) Karapetis 2003 (n=15) : Se 100%, Sp 72%	See above
Systematic reviews				
Bourguet 2007	Patients with testicular cancer	**FDG-PET**	No new primary studies found since the previous FNCLCC report (FNCLCC 2003).	See above

Evaluation of residual mass

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with testicular cancer	**FDG-PET**	One primary study identified. Becherer 2005 (n=48): per-lesion analysis.	See above
Systematic review				
Bourguet 2007	Patients with testicular cancer	**FDG-PET**	Two new primary studies (Becherer 2005, Pfannenberg 2004) identified since previous FNCLCC report (in total 76 patients included). Better sensitivity (80% vs. 73%) and specificity (100% vs. 73%, p<0.001) for PET compared to CT in one study (for diagnosis of viability of residual masses; per-lesion analysis), but comparable sensitivity (62% vs. 62%) and specificity (83% vs. 72%) for PET and CT/MRI in other study.	See above

Primary studies

	Population	Index test	Outcome	Results	Comments
Oechsle 2008	Patients with NSGCT and a primary or metastatic retroperitoneal tumour of at least 5 cm or with distant metastases at the time of primary diagnosis or first relapse (n=121)	**FDG-PET** Reference standard: histology (resection specimen)	Evaluation of residual mass	Sensitivity 70% (95%CI 58-81%) Specificity 39% (95%CI 26-53%) PPV 59% (vs. 55% for CT) NPV 51%	Prospective study No information on blinding Consecutive patients?

PROSTATE CANCER

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with prostate cancer	**FDG-PET** **FDG-PET/CT**	Four primary studies identified (Jadvar 2003, Schoder 2005, Oyama 2003, Chang 2003), with different objectives (mainly detection and/or staging of recurrent disease) and different reference standards. Chang 2003 (n=24, recurrence): Se 75% (48-93%), Sp 100% (63-100%) Jadvar 2003 (n=12, recurrence): Se 50% (16-84%), Sp 75% (19-99%) Oyama 2003 (n=48, recurrence): partial verification Schoder 2005 (n=91, recurrence) : Se 32% (22-43%), Sp 0% (0-71%)	Good-quality HTA *Search date*: 2003 - March 2008 *Databases*: Medline, EMBASE, Central, Scopus *Meta-analysis* using the DerSimonian and Laird random effects method. Software: RevMan software version 5.0.
Systematic reviews				
Bourguet 2006	Patients with prostate cancer	**FDG-PET**	Four new primary studies found (Oyama 2002, Chang 2003, Schoder 2005, Morris 2005).	See above.

	Population	Index test	Results	Comments
			Conclusions: Diagnosis (no change): no indication. Locoregional staging: no indication. Recurrence (no change): to be confirmed in peer-reviewed protocols: FDG-PET can be useful for the detection of local recurrence and occult disease after radical treatment. Evaluation of treatment: no indication.	

BLADDER CANCER

	Population	Index test	Results	Comments
HTA reports				
AHRQ 2008	Patients with bladder cancer	**FDG-PET** **FDG-PET/CT**	Staging of primary tumour: Two prospective studies (Drieskens 2005, Liu 2003) including a total of 88 patients. Index test: FDG-PET. Reference standard: Histology/biopsy (or clinical follow-up in Drieskens 2005). *Sensitivity*: 53 – 77% *Specificity*: 72 – 94% Considerable heterogeneity. Staging of recurrent disease: One retrospective study identified (Jadvar 2008), including 35 patients. Index test: both FDG-PET (n=17) and FDG-PET/CT (n=18). Change in management in 17%.	See above.

BRAIN CANCER

Diagnosis

	Population	Index test	Results	Comments
HTA				
AHRQ 2008	30 patients undergoing evaluation for brain tumours (both primary diagnosis and recurrences)	**FDG-PET** Reference standard: Histology/biopsy or clinical follow-up (mean 20 mo)	Chen 2006 (n=30) FDG-PET Sensitivity= 61% (95%CI 39%-79%) Specificity= 43% (95%CI 12%-80%) FDOPA-PET Sensitivity= 96% (95%CI 76%-99%) Specificity= 43% (95%CI 12%-80%) FDOPA PET is particularly useful for imaging of low-grade	Prospective study

			tumours and evaluating recurrent tumours; its sensitivity is higher than that of FDG PET for the same purposes. However, its specificity is low.	

Primary studies

	Population	**Index test**	**Outcome**	**Results**	**Comments**
Pöpperl 2007	54 adult patients (mean age: 49±17 yrs, range: 18–76 yrs) with suspected supratentorial primary gliomas	**FET-PET** Comparator: MRI Reference standards: Histopathology (n=21) or clinical follow-up (n=17)	Grade of malignancy: low grade vs high grade	Histopathological analyses revealed glioma tissue in all patients, 46/54 suffering from astrocytic tumours and 8/54 from mixed oligoastrocytomas Differentiation between LG and HG with FET-PET Se: 93% (95% CI 76%-99%) Sp: 100% (95%CI 80%-100%) Conclusion : histopathologic examinations remain the gold standard for establishing tumour grade; however, dynamic FET uptake evaluations contribute significantly to predicting ultimate histological findings	Study conducted in Germany. Prospective or retrospective design: not clear
Roessler 2007	27 patients (mean age 42 years, range 11-77 years) with suspected cerebral gliomas	**MET-PET** Reference standard: histopathology	Detection and identification (grading) of brain tumour	All patients had a brain tumour: anaplastic glioma or anaplastic mixed glioma (WHO grade III) or glioblastoma (WHO grade IV) in 11 patients, low-grade astrocytoma or mixed glioma (WHO grade II) in 8 patients, pure oligodendroglioma (WHO grade II or III) in 8 patients MET-PET diagnosed 26 tumours: Se 96% (95% CI: 79% - 99%)	Prospective study

Staging

	Population	Index test	Results	Comments
HTA				
AHRQ 2008	Patients with suspected primary glioma (Cher 2006, Stockhammer 2007) or patients with primary astrocytomas (Liu 2006)	**FDG-PET** Reference standard: Histology/biopsy	Cher 2006 (n=16 patients) Se: 63% (95%CI 35%-85 %) Specificity (not calculated) Liu 2006 (n=26 patients) Se: 63% (95%CI 38%-84%) Sp: 100% (95%CI 59%-100%) Stockhammer 2007 (n=25 patients) Sensitivity= 75% (95%CI 47%-92%) Specificity= 0% (95%CI 0%-37%)	Prospective studies (Cher and Liu); retrospective study (Stockhammer)

Detection of recurrence

	Population	Index test	Results	Comments
HTA				
AHRQ 2008	28 patients with glioblastoma multiforme after surgical and/or conservative treatment (mean follow-up: 13 mo)	**FDG-PET** Reference standard: MRI and MET-PET (carbon-11 methionine and positron emission tomography)	Potzi 2007 (n=28 patients) Detection of recurrences: Sensitivity= 11% (95%CI 2%-36%) Specificity= not defined (0/0) FDG-PET vs. survival > 12 mo: Sensitivity= 7% (95%CI 0.4%-38%) Specificity= 14% (95%CI 2.5%-44%) FDG PET is of limited value in the work-up of recurrent GBM because of its low sensitivity and the fact that it allows no prediction of the outcome	Retrospective study Reference standards are index tests

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
McCarthy 2009	38 patients referred with possible brain tumour recurrence (n = 32), or newly	**FDG-PET** Reference standards: Histopathology (n=21) or	Diagnosis of the presence or absence of brain tumour	In the whole sample of 38 patients, there were 27 brain tumours (including 23 gliomas) and 11 non-tumorous lesions.	Retrospective study Possibility of review bias (unclear)

	Population	Index test	Outcome	Results	Comments
	diagnosed with brain tumour (n=6)	clinical follow-up (n=17)		Of the 21 patients with histological results, 18 were shown to have tumours and 3 were shown to have non-tumorous lesions FDG-PET vs both reference standards (n=38): Se: 74% (95% CI 53%-83%) Sp: 73% (95%CI 39%-86%) FDG-PET vs histopathology (n=21) : Se 72% (95%CI 47%-89%) Sp 33% (95%CI 2%-87%)	
Pöpperl 2006	24 patients with a mean age of 49±14 years and histopathologically proven malignant gliomas (5 anaplastic astrocytomas, 19 glioblastomas)	**FET-PET** Reference standards: biopsy or clinical follow-up (up to 87 months) Comparator: MRI	Detection of recurrence after locoregional radio-immunotherapy (RIT) Survival (in months)	Among the 17 of 24 patients who presented with tumour progression, 10 had tumour recurrence and 7 had re-growth of residual tumour FET-PET: Se 94% (95%CI 69%-99%) Sp 71% (95%CI 30%-95%) PET scan demonstrated the highest discrimination capacity between patients with tumour recurrence and tumour-free patients at a threshold value of 2.4 for the TU_{max}/BG ratio. Survival times were significantly longer ($p<0.05$) in patients presenting with values below this threshold. Conclusion: FET-PET is a sensitive tool for monitoring the effects of high local radiation doses given by intracavitary RIT. Focally increased FET uptake is an early and reliable indicator of tumour progression.	TU_{max} : maximal tumoral uptake BG: Background Limit: lack of histological confirmation in 15/24 patients, and especially in 5/7 patients who were considered tumour free. Kaplan-Meier curves were provided for survival
Rachinger 2005	45 consecutive	**FET-PET**	Detection of tumour	In patients who already had suspected	Retrospective study

	Population	Index test	Outcome	Results	Comments
	patients with gliomas including patients having suspected tumor recurrence or progression by follow-up MRI (n=36) and patients treated by RIT (n=9).	Reference standards: histopathology or clinical follow-up Comparator: MRI	recurrence or tumour progression after treatment	recurrent tumour on the basis of MRI, FET-PET yielded: Se 100% (95%CI 86%-100%) Sp 93% (95%CI 64%-100%) FET-PET is useful to differentiate side effects of therapy from tumour recurrence.	including all patients with suspected tumour recurrence revealed by MRI

Prognosis

Primary studies

	Population	Index test	Outcome	Results	Comments
FDG-PET					
Spence 2008	22 patients with glioblastoma multiforme (GBM) with a median age of 56 years before radiotherapy	**FMISO PET** Reference standard: follow-up (survival) and MRI criteria (progression)	Overall survival and time to progression (TTP) in months	Multivariate analyses for survival and TTP against the covariates HV (or T/B max), MRI T1Gd volume, age, and KPS reached significance only for HV (or T/B max; $p<0.03$). The multivariate Cox model indicates that an increase in HV on the order of 7 to 8 cm^3 is associated with a 50% reduction in survival or TTP. An increase in T/B max of 0.25 (or 0.35 for TTP) is associated with a 50% reduction in survival (or TTP). Conclusion: FMISO PET is useful to describe the volume and intensity of hypoxia in GBM before radiotherapy, which are strongly associated with poorer TTP and survival. This type of imaging could be integrated into treatment strategies to target hypoxia more aggressively in GBM.	Survival and TTP were calculated from the date of surgery and the most recent follow-up information. All analyses were completed with standard censoring procedures for survival analysis. The prognostic variables considered were FMISO hypoxic volume (HV) and tissue to blood concentration (T/B) max; age; Karnofsky performance status (KPS); extent of resection; and the MRI volumes, TO, TI Gd, T2, and T2-TO.
Ceyssens 2006	52 patients (mean age, 41.5 years; range, 3-72	**MET-PET**	Overall median survival	Overall median survival was 34.9 months.	Retrospective study

	Population	Index test	Outcome	Results	Comments
	years)	Reference standard: clinical follow-up		In a proportional hazard Cox regression model with age, WHO grading, and MET uptake index, only WHO grading was significantly predictive of survival (p =0.015), whereas age and MET uptake index were not significantly predictive. Differences in survival were found only for WHO low-grade versus high-grade (III-IV) tumours (p=0.017). No thresholds could be found at which MET could be considered predictive of survival (Kaplan-Meier statistics).	Follow-up: until the last clinical contact with normal findings (n = 27; average, 24.2 months after PET) or until death (n = 19; average, 13,5 months after PET) 6 patients were lost to follow-up Survival curves (Kaplan-Meier) and proportional hazard Cox regression model.
Van Laere 2005	30 patients (age 40.4± 15.6 years), on average 4.0 years after therapy for a primary brain tumour (23 grade II–IV astrocytomas, four oligodendrogliomas and three mixed oligo-astrocytomas)	**MET-PET** **FDG-PET** Reference standards: histopathology or clinical follow-up or radiological imaging	Overall median survival	Overall median survival was 15.0 months. MET showed pathologically increased uptake in 28/30 scans and FDG in 17/30. Kaplan-Meier survival analysis: significant differences were found for both FDG (cut-off 0.8, p=0.007) and MET (cut-off 2.2, p=0.014). MET alone was the best prognostic predictor in the subgroup of patients with primary astrocytoma (n=23). Conclusion: FDG and MET-PET studies provide complementary prognostic information in patients with suspected brain tumour recurrence or progression after primary therapy.	Retrospective study Minimum follow-up of 1 year or until death

CARDIOVASCULAR DISEASE

Myocardial perfusion evaluation

	Population	Index test	Results	Comments
Systematic reviews				
Nandalur 2008	Patients with suspected coronary artery disease	**PET (mixed tracers)** Standard: Catheter x-ray angiography Objective: Diagnosis of coronary artery disease (≥50% diameter stenosis as threshold for significant CAD)	Nineteen primary studies included, involving 1442 patients. Overall low quality of the included studies. Sensitivity: 0.92 (95%CI 0.90–0.94) (no heterogeneity) Specificity: 0.85 (95%CI 0.79–0.90) (significant heterogeneity)	Good-quality systematic review *Search date*: July 2007 *Databases*: Medline, EMBASE, handsearching No language restriction Meta-analysis performed
Beanlands 2007a	Patients with suspected coronary artery disease	**PET (mixed tracers)** Standard: Catheter x-ray angiography in majority of studies (mixed definitions of significant CAD) Objective: Diagnosis of coronary artery disease	Fourteen primary studies included on myocardial perfusion evaluation (13 of them included in Nandalur 2008). PET: Mean sensitivity: 89% (range 83-100%) Mean specificity: 89% (range 73-100%) 16-slice MDCT: Mean sensitivity: 98% (range 85-100%) Mean specificity: 86% (range 67-98%) MR angiography: (1.5 T) Mean sensitivity: 72% (range 38-90%) Mean specificity: 87% (range 73-100%) Dobutamine stress echo: Mean sensitivity: 90% (range 86-96%) Mean specificity: 84% (range 80-86%)	Moderate-quality systematic review *Search date*: June 2005 *Databases*: Medline, EMBASE, CDSR, AHRQ website No meta-analysis

Myocardial viability

	Population	Index test	Results	Comments
HTA reports				
MAS 2005	Patients with ischemic heart disease and left ventricular dysfunction	**PET**	Update of ICES HTA 2001, which was included in the previous KCE report. Nine new primary studies identified (Koch 2001; Tani 2001; Wiggers 2001; Lund 2002; Nowak 2003; Wiggers 2003; Korosoglou 2004; Barrington 2004; Schmidt 2004). PET: Median sensitivity: - Mean LVEF ≤35%: 90% (range 75-100%) - Mean LVEF >35%: 89% (range 76-100%) Median specificity: - Mean LVEF ≤35%: 67% (range 33-81%) - Mean LVEF >35%: 85% (range 35-91%) Comparison with other tests: - Observational studies suggest that FDG-PET has the highest sensitivity but dobutamine echocardiography has the highest specificity for predicting regional LV function recovery after revascularization. - FDG-PET and dobutamine echocardiography appear to have comparable diagnostic accuracy. - Thallium SPECT appears to be inferior to PET and dobutamine echocardiography for predicting regional function recovery. It has been shown to underestimate viability in patients with severe LV dysfunction (<25%). - FDG-PET detected viable myocardium in 43% to 50% of patients found to have non-viable myocardium by thallium-201 SPECT. - FDG-SPECT appears to have good overall concordance with FDG-PET in detecting viable myocardium; however, it may overestimate viability in severely dysfunctional regions or in regions with severely reduced FDG uptake on PET.	Moderate-quality HTA *Search date*: April 2005 *Databases*: Medline, EMBASE, Cochrane Library, INAHTA database Meta-analysis performed where possible
Systematic reviews				
Beanlands 2007a	Patients with ischemic heart disease and left	**PET**	Eight primary studies on myocardial viability evaluation found in addition to the systematic review of Bax 2001.	See above

	Population	Index test	Results	Comments
	ventricular dysfunction		PET: Mean sensitivity: 91% (range 80-100%) Mean specificity: 61% (range 44-92%) Dobutamine stress MR: Mean sensitivity: 91% (range 77-100%) Mean specificity: 94% (range 69-100%) Late Gadolinium enhancement MR: Mean sensitivity: 81% (range 64-99%) Mean specificity: 83% (range 72-98%)	

Primary studies

	Population	Index test	Outcome	Results	Comments
Beanlands 2007b	Patients with severe left ventricular (LV) dysfunction and suspected coronary disease being considered for revascularization, heart failure, or transplantation work-ups or in whom PET was considered potentially useful (n=430)	Management assisted by FDG-PET (n=218) or standard care (n=212)	Composite of cardiac death, myocardial infarction, or recurrent hospital stay for cardiac cause, within 1 year	At 1 year, the cumulative proportion of patients who had experienced the composite event was 30% (PET arm) vs. 36% (standard arm) (RR 0.82, 95%CI 0.59-1.14; p=0.16). HR for the composite outcome, PET vs. standard care: 0.78 (95%CI 0.58-1.1; p=0.15).	RCT. Blinded study. Intention-to-treat-analysis.

Prognosis

	Population	Index test	Results	Comments
Systematic reviews				
Beanlands 2007a	Patients with coronary artery disease	**82Rb-PET**	Five primary studies identified on prognosis (Yoshinaga 2006, Chow 2006, Marwich 1997, Marwick 1995, MacIntrye 1993). Hard cardiac event rates: 0.09-0.9% for normal PET results vs. 7% for abnormal PET results.	See above.

Primary studies

	Population	Index test	Outcome	Results	Comments
Tio 2009	Patients with advanced ischemic heart disease (n=480)	**FDG-PET**	Cardiac death	After controlling for age and sex, the following parameters were associated with cardiac death: myocardial perfusion reserve (MPR) measured by FDG-PET, family history, previous	Retrospective analysis of prospective database. Seventeen patients excluded from analysis because gating not

	Population	Index test	Outcome	Results	Comments
				myocardial infarction, LVEF, left ventricular end-diastolic volume, aspirin, diuretics, and digoxin. MPR was associated with a HR for cardiac death of 4.11 (95%CI 2.98–5.67) per SD decrease, whereas the risk for LVEF was 2.76 (2.00–3.82) per SD decrease.	possible. Another 119 patients excluded because of undergoing a PET-driven revascularisation. Mean follow-up among survivors: 85 months (range 1-138 months).
Santana 2008	Patients with ischemic cardiomyopathy (n=104)	**82RB/gated FDG-PET**	Cardiac death	Using univariate analysis, none of the variables (including PET) were predictive of cardiac death.	Unclear if prospective study. Consecutive patients. Mean follow-up: 21.6 +/- 14 months (range 0.23-54 months).

INFECTIOUS DISEASES

Osteomyelitis

	Population	Index test	Results	Comments
HTA reports				
CADTH 2008	Patients with suspected osteomyelitis	**FDG-PET**	One systematic review identified (Termaat 2005, cfr. infra). Three additional primary studies found: 1. Basu 2007: • Indication: differentiation of Charcot's neuroarthropathy from osteomyelitis and soft-tissue infection (n=63) • Comparator: MRI • Results: sensitivity 100% (vs. 77% for MRI) 2. Schwegler 2008: • Indication: diabetic foot ulcer and osteomyelitis • Comparators: MRI, 99mTc-MOAB • Results: sensitivity 29% (vs. 86% for MRI), specificity 92% (vs. 92% for MRI) 3. Hakim 2006: • Indication: chronic osteomyelitis of the mandible • Comparator: bone scintigraphy with SPECT • Results: sensitivity 64% (vs. 84% for SPECT), specificity 78% (vs. 33% for SPECT)	Moderate-quality HTA *Search date*: March 2008 Databases: Medline, EMBASE, CRD database, Cochrane Library, Google Restriction to English No meta-analysis performed
Systematic reviews				
Termaat	Patients with	**FDG-PET**	Four (low-quality) primary studies on FDG-PET included.	Good-quality SR

	Population	Index test	Results	Comments
2005	suspected chronic osteomyelitis		Pooled sensitivity of FDG-PET was significantly higher than that of other tests (p<0.05), estimated at 96% (95% CI 88-99%). Pooled specificity of FDG-PET (91%; 95%CI 81-95%) was significantly higher than that of leukocyte scintigraphy, bone scintigraphy, and MRI, but not significantly different from combined bone and leukocyte scintigraphy, and combined bone and gallium scintigraphy.	*Search date*: July 2003 *Databases*: Medline, EMBASE, Current Contents Restriction to English Meta-analysis performed

Prosthetic joint infections

	Population	Index test	Results	Comments
HTA reports				
CADTH 2008	Patients with suspected prosthetic hip or knee joint infection	**FDG-PET**	One low-quality meta-analysis identified (Prandini 2006), which included 6 primary studies on PET scan. Sensitivity values for the imaging test methods ranged from 70.1% (Gallium) to 95.2% (polyclonal human-immune globulin), with FDG-PET sensitivity being 94.1%. Specificity estimates ranged from 75.2% (bone scan) to 89.1% (scintigraphy with white blood cells), with FDG-PET specificity being 87.3%. FDG-PET had the highest accuracy (91.9%) of all methods. One additional primary study was found (Pill 2006), which was also included in Kwee 2008.	See above
Systematic reviews				
Kwee 2008	Patients with suspected prosthetic hip or knee joint infection	**FDG-PET**	Inclusion of 11 primary studies. Pooled sensitivity: 82.1% (95%CI 68.0-90.8%) Pooled specificity: 86.6% (95%CI 79.7-91.4%) Heterogeneity among the results of individual studies was present (I^2 = 68.8%).	Good-quality systematic review *Search date*: May 2008 *Databases*: Medline, EMBASE, handsearching No language restriction Meta-analysis performed

Fever of unknown origin

	Population	Index test	Results	Comments
HTA reports				
CADTH 2008	Patients with fever of unknown origin	**FDG-PET**	One primary study identified (Bleeker-Rovers 2007), involving 70 patients (of which 43 had both PET and CT). Overall (n=70): sensitivity 88%, specificity 77%. Comparison PET vs. CT (n=43): PPV 65% vs. 48%, NPV 90% vs. 86%.	See above

Primary studies

	Population	Index test	Outcome	Results	Comments
Keidar 2008	Patients with fever of unknown origin (n=48)	**FDG-PET/CT** Standard: Histopathology, microbiology/ serology, clinical diagnostic criteria defined by the treating physician, clinical and imaging follow-up	Detection of infection focus	*Sensitivity*: 100% (22/22; 95%CI 85-100%) *Specificity*: 81% (21/26; 95%CI 61-93%)	Prospective study. Consecutive patients. Differential verification. Unclear if PET/CT was part of follow-up.

Infections of the vertebral column

	Population	Index test	Results	Comments
HTA reports				
CADTH 2008	Patients with suspected infection of the vertebral column	**FDG-PET**	One low-quality meta-analysis identified (Prandini 2006). FDG-PET had the highest sensitivity and accuracy (100% and 90%, respectively) of all the test methods considered.	See above

Vascular infections

Primary studies

	Population	Index test	Outcome	Results	Comments
Keidar 2007	Patients with a suspected prosthetic vascular graft infection (clinical signs) (n=39)	**FDG-PET/CT** Standard: Histopathology, microbiology, imaging/clinical follow-up	Diagnosis of vascular graft infection.	*Sensitivity*: 93% (14/15; 68-100%) *Specificity*: 92% (22/24; 73-99%)	Prospective study. Consecutive patients. Differential verification. Unclear if PET/CT was part of follow-up.

EPILEPSY SURGERY

Prognosis

	Population	Index test	Results	Comments
HTA reports				
Whiting P-NHS 2006 (UK)	All adults and children with refractory epilepsy being considered for epilepsy surgery	**PET**: FDG-PET in 19 studies, 4 studies used additionally other tracers. PET compared to conventional MRI, volumetric MRI, ictal SPECT, interictal SPECT. Reference standards: ictal EEG or multimodal evaluation (mostly ictal EEG, neuropsychological testing, Wada testing and other imaging techniques) or site of surgery	19 primary studies on diagnostic accuracy included 7 studies on prognosis (surgery outcome) PET compared to conventional MRI, volumetric MRI, ictal SPECT, interictal SPECT because most of the primary studies did not provide data on PET results for patients in whom the reference standard failed to localise a seizure focus, no 2X2 tables were constructed and sensibility /specificity were not reported authors' conclusion: results do little to inform clinical practice, owing to the limitations of the included studies	Good quality. *Databases*: Medline, Embase, Biosis, Pascal, Science citation index, Lilacs, no language restriction. *Search period* (until Dec 2003) already covered in literature (up to June 2004) included in KCE report 22 *No meta-analysis* *Comments:* -selection bias in some studies: only patients already selected to undergo surgery or with a good surgery outcome -reference standard biased: ictal EEG and multimodal evaluation may fail to find the epilepsy focus; side of surgery is only available in patients undergoing surgery; results of index test might have been used in decision to perform surgery -possible heterogeneity in exact use of technology and the way it was applied; or in result interpretation (dependent on skills of evaluator)

Primary studies

	Population	Index test	Outcome	Results	Comments
Yun CH 06	N=193 consecutive patients; N for PET = 179- N for SPECT/SISCOM =136- Patients with refractory neocortical focal epilepsy (frontal, temporal, parietal, occipital, multifocal) included- Adults (26+-7 years)- All patients undergo surgery	**FDG-PET** Reference standard: Site of surgery as decided based on MRI and EEG	Value of PET (by visual and by SPM analysis) in localizing epileptogenic focus as compared to gold standard is low. -Surgical outcome: 58% (111/193 patients) seizure free. -focal MRI lesion, focal ictal scalp EEG and FDG-PET	-67 of 107 (63%) patients seizure free correctly localised by PET (42% by ictal SPECT) -29 of 72 (40%) patients not seizure free correctly localised by PET	-Low quality study-Retrospective study -large cohort size -no information on blinding - Spectrum of patients is not representative: only patients that underwent surgery (incorporation bias or selection bias)

	Population	Index test	Outcome	Results	Comments
	(focal surgery only)- visible lesion on MRI is no exclusion criterion		significant independent predictors of post-surgical seizure-free outcome (univariate analysis and multiple logistic regression analysis): OR (PET) 2.49 (95%CI: 1.35-4.61; p= 0.004)	(36% by ictal SPECT)	-result of surgery can be dependent on other variables (e.g. eloquent cortex not resected)
Gaillard W 07	N=38 -children (5.8 yrs, range 0.9-11.9) taking anti-epileptic drugs -at least 3 complex partial seizures at study entrance (range 3-200) -no obvious etiology (trauma, infection...) or benign (rolandic) epilepsy -no gross structural or mass lesion on MRI (blinded) -mean epilepsy duration 1.1 year (0.3-2.3)	**FDG-PET**: -Average of 3.4 FDG PET scan per child (over 3.0 years +-1.3 yrs) -abnormal PET: automatically calculated asymmetry index beyond 2SD of mean	Initial normal PET significantly more likely to remain in good seizure control (p<0.01) Model combining MRI and PET strongly predictive of clinical course. If normal MRI and PET scan initially then more likely to remain in good seizure control over 3 to 4 years. No evidence for progression of hypometabolism or development of bilateral hypometabolism	Logistic regression: Normal MRI higher predictive value (OR 0.036 range 0.004-0.332 for poor outcome; p<0.01) than normal PET (OR 0.215 range 0.023-2.04; p<0.20)	Moderate quality study- Prospective blinded cohort study -small cohort sample -clinical data collected at outpatient clinic -no validation in an independent group ("test set") of patients

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
Whiting P- NHS 2006 (UK)	See above.			

PARKINSON

Diagnosis

	Population	Index test	Results	Comments
Systematic reviews				
Suchowersky O (American Academy of Neurology) 2006	Idiopathic Parkinson Disease (PD) and atypical parkinsonian syndromes (APS)	**PET** (all radioligands)	No publications of sufficient quality for prognosis--- 1 case control study by Antonini A 1998: 56 PD and 48 APS patients studied by FDG PET compared to clinical diagnosis as gold standard. Conclusion of SR: insufficient evidence to support or refute PET (all radioligands) to make the differential diagnosis between PD and APS	Good quality systematic review

ALZHEIMER'S DEMENTIA (AD) AND NON-AD DEMENTIA

Diagnosis

	Population	Index test	Results	Comments
HTA reports				
SBU 2008	Alzheimer's dementia (AD) versus non-Alzheimer dementia. Diagnosis of AD based on internationally accepted standardised clinical criteria and/or post-mortem histopathology	**FGD-PET**	Reference test: clinical diagnosis based on internationally accepted standardised clinical criteria: Azari 1993: 19 probable AD (retrosp. case control study)- Herholz 2002: 395 probable AD (retrospective case control study)- Smith 1992: 45 AD (prospective case control study) Reference test: histopathology: Hoffman 2000: 22 possible AD (retrospective cohort study)- Silverman 2001: 120 possible AD (retrospective cohort study)- Silverman 2003: 167 possible progressive dementia patients (prospective cohort study)	Sensitivity as well as specificity >0.8 and LR+>5: Azari 1993; Herholz 2002;Smith 1992 Hoffman 2000: sensitivity 0.93; specificity 0.63 and LR+ 2.5 Silverman 2001 and Silverman 2003: sensitivity 0.94-0.95; specificity 0.73-0.79; LR+ 3.5-5.0 Conclusion systematic review: FDG PET scan has moderate value (Evidence Grade 2) differentiating AD from normal subjects and from other dementia disorders.
Systematic reviews				
Yuan Y 2009	Patients with mild cognitive impairment: prediction of conversion to Alzheimer's disease during follow-up. Diagnosis of AD based on internationally accepted standardised clinical criteria and/or post-mortem histopathology	**FDG-PET** Comparators: SPECT, structural MRI	24 retrospective studies included (6 FDG PET;8 SPECT;10 MR imaging) Weighted summaries: Sensitivity: FDG-PET 89% (82-94%); SPECT 64% (77-89%); MRI 73% (65-80%) Specificity:FDG-PET 57% (78-90%);SPECT 70% (63-77%); MRI 81% (76-85%) LR+: FDG-PET 4,6 (3,2-6,7);SPECT 2,6 (1,4-4,6); MRI 3,5 (2,6-4,6) LR-: FDG-PET 0,15 (0,05-0,48);SPECT 0,32 (0,21-0,49); MRI 0,37 (0,29-0,48) OR: FDG-PET 40,1 (18,5-69,7);SPECT 9,3 (4,5-19,3); MRI 10,6 (6,6-17,0) No significant difference (p>0.05) for sensitivity, specificity and LR- between all techniques. Significant (p<0.05) better LR+ and OR for FDG-PET.	Good-quality meta-analysis Conclusions by the authors: FDG-PET performs slightly better than SPECT and structural MRI in the prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment and parallel performance was found between SPECT and MR imaging. Cave: Heterogeneity highly significant for LR- (FDG-PET) and LR+ (SPECT); meta-regression no clear explaning factors Marked asymmetry suggesting publication bias.

11 REFERENCES

1. Cleemput I, Dargent G, Poelmans J, Camberlin C, Van den Bruel A, Ramaekers D. HTA positron emission tomography imaging in Belgium. Brussels: Belgian Health Care Knowledge Centre (KCE). 2005;KCE Reports vol. 22C. Ref. D2005/10.273/31.
2. ZonMw. PET gepast gebruik(t). Den Haag: ZonMw; 2007.
3. Medical Services Advisory Committee. Positron emission tomography for recurrent colorectal cancer. Canberra: MSAC; 2008.
4. Kumar R, Dhanpathi H, Basu S, Rubello D, Fanti S, Alavi A. Oncologic PET tracers beyond [18F]FDG and the novel quantitative approaches in PET imaging. Q J Nucl Med Mol Imaging. 2008;52:50-65.
5. Piccini P, Whone A. Functional brain imaging in the differential diagnosis of Parkinson's disease. Lancet neurol. 2004;3(5):284-90.
6. Shih M, Hoexter MQ, Andrade LAFd, Bressan RA. Parkinson's disease and dopamine transporter neuroimaging - a critical review. Sao Paulo Med J. 2006;124(3):168-75.
7. Schreckenberger M, Hagele S, Siessmeier T, Buchholz H-G, Armbrust-Henrich H, Rosch F, et al. The dopamine D2 receptor ligand 18F-desmethoxyfallypride: an appropriate fluorinated PET tracer for the differential diagnosis of parkinsonism. Eur J Nucl Med Mol Imaging. 2004;31(8):1128-35.
8. Knottnerus JA, van Weel C, Muris JWM. Evaluation of diagnostic procedures. BMJ. 2002;324(7335):477-80.
9. Ledley RS, Lusted LB. Reasoning foundations of medical diagnosis; symbolic logic, probability, and value theory aid our understanding of how physicians reason. Science. 1959;130(3366):9-21.
10. Guyatt GH, Tugwell PX, Feeny DH, Haynes RB, Drummond M. A framework for clinical evaluation of diagnostic technologies. CMAJ. 1986;134(6):587-94.
11. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making. 1991;11(2):88-94.
12. Sackett DL, Haynes RB. The architecture of diagnostic research. BMJ. 2002;324(7336):539-41.
13. Bossuyt PM, Irwig L, Craig J, Glasziou P. Comparative accuracy: assessing new tests against existing diagnostic pathways. BMJ. 2006;332(7549):1089-92.
14. Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ. 2008;336(7653):1106-10.
15. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999;282(11):1061-6.
16. Rutjes AWS, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PMM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ. 2006;174(4):469-76.
17. Bossuyt PM, Lijmer JG, Mol BW. Randomised comparisons of medical tests: sometimes invalid, not always efficient. Lancet. 2000;356(9244):1844-7.
18. Biesheuvel CJ, Grobbee DE, Moons KGM. Distraction from randomization in diagnostic research. Ann Epidemiol. 2006;16(7):540-4.
19. Lord SJ, Irwig L, Simes RJ. When is measuring sensitivity and specificity sufficient to evaluate a diagnostic test, and when do we need randomized trials? Ann Intern Med. 2006;144(11):850-5.
20. National Collaborating Centre for Acute Care, February 2005. Diagnosis and treatment of lung cancer. National Collaborating Centre for Acute Care, London. Available from www.rcseng.ac.uk.
21. Gould MK, Fletcher J, Iannettoni MD, Lynch WR, Midthun DE, Naidich DP, et al. Evaluation of patients with pulmonary nodules: when is it lung cancer? ACCP evidence-based clinical practice guidelines (2nd edition). Chest. 2007;132(3 Suppl):108S-30S.
22. Facey K, Bradbury I, Laking G, Payne E. Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers. Health Technology Assessment. 2007;11(44):iii-iv, xi-267.

23. Tsunezuka Y, Shimizu Y, Tanaka N, Takayanagi T, Kawano M. Positron emission tomography in relation to Noguchi's classification for diagnosis of peripheral non-small-cell lung cancer 2 cm or less in size. World Journal of Surgery. 2007;31(2):314-7.
24. Gould MK, Ananth L, Barnett PG, Veterans Affairs SCSG. A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules. Chest. 2007;131(2):383-8.
25. Silvestri GA, Gould MK, Margolis ML, Tanoue LT, McCrory D, Toloza E, et al. Noninvasive staging of non-small cell lung cancer: ACCP evidenced-based clinical practice guidelines (2nd edition). Chest. 2007;132(3 Suppl):178S-201S.
26. Wahidi MM, Govert JA, Goudar RK, Gould MK, McCrory DC. Evidence for the treatment of patients with pulmonary nodules: when is it lung cancer? ACCP Evidence-Based Clinical Practice Guidelines (2nd edition). Chest. 2007;132(3):94S-107S.
27. Herder GJ, van Tinteren H, Golding RP, Kostense PJ, Comans EF, Smit EF, et al. Clinical prediction model to characterize pulmonary nodules: validation and added value of 18F-fluorodeoxyglucose positron emission tomography. Chest. 2005;128(4):2490-6.
28. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. Arch Intern Med. 1997;157(8):849-55.
29. Cronin P, Dwamena BA, Kelly AM, Carlos RC. Solitary pulmonary nodules: meta-analytic comparison of cross-sectional imaging modalities for diagnosis of malignancy. Radiology. 2008;246(3):772-82.
30. Kaira K, Oriuchi N, Shimizu K, Ishikita T, Higuchi T, Imai H, et al. Evaluation of thoracic tumors with (18)F-FMT and (18)F-FDG PET-CT: a clinicopathological study. International Journal of Cancer. 2009;124(5):1152-60.
31. Buck AK, Hetzel M, Schirrmeister H, Halter G, Moller P, Kratochwil C, et al. Clinical relevance of imaging proliferative activity in lung nodules. European Journal of Nuclear Medicine & Molecular Imaging. 2005;32(5):525-33.
32. Hashimoto Y, Tsujikawa T, Kondo C, Maki M, Momose M, Nagai A, et al. Accuracy of PET for diagnosis of solid pulmonary lesions with 18F-FDG uptake below the standardized uptake value of 2.5. Journal of Nuclear Medicine. 2006;47(3):426-31.
33. Ferran N, Ricart Y, Lopez M, Martinez-Ballarin I, Roca M, Gamez C, et al. Characterization of radiologically indeterminate lung lesions: 99mTc-depreotide SPECT versus 18F-FDG PET. Nuclear Medicine Communications. 2006;27(6):507-14.
34. Wang F, Wang Z, Yao W, Xie H, Xu J, Tian L. Role of 99mTc-octreotide acetate scintigraphy in suspected lung cancer compared with 18F-FDG dual-head coincidence imaging. Journal of Nuclear Medicine. 2007;48(9):1442-8.
35. van Tinteren H, Hoekstra OS, Smit EF, van den Bergh JHAM, Schreurs AJM, Stallaert RALM, et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. Lancet. 2002;359(9315):1388-93.
36. Ohno Y, Koyama H, Onishi Y, Takenaka D, Nogami M, Yoshikawa T, et al. Non-small cell lung cancer: whole-body MR examination for M-stage assessment--utility for whole-body diffusion-weighted imaging compared with integrated FDG PET/CT. Radiology. 2008;248(2):643-54.
37. Ohno Y, Koyama H, Nogami M, Takenaka D, Yoshikawa T, Yoshimura M, et al. STIR turbo SE MR imaging vs. coregistered FDG-PET/CT: quantitative and qualitative assessment of N-stage in non-small-cell lung cancer patients. Journal of Magnetic Resonance Imaging. 2007;26(4):1071-80.
38. Quaia E, Tona G, Gelain F, Lubin E, Pizzolato R, Boscolo E, et al. Integrated fluorine-18 fluorodeoxyglucose (18F-FDG) PET/CT compared to standard contrast-enhanced CT for characterization and staging of pulmonary tumors eligible for surgical resection. Acta Radiologica. 2008;49(9):995-1004.
39. Chin K, Finger PT, Kurli M, Tena LB, Reddy S. Second cancers discovered by (18)FDG PET/CT imaging for choroidal melanoma. Optometry. 2007;78(8):396-401.

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

41. Kim YK, Lee KS, Kim B-T, Choi JY, Kim H, Kwon OJ, et al. Mediastinal nodal staging of nonsmall cell lung cancer using integrated 18F-FDG PET/CT in a tuberculosis-endemic country: diagnostic efficacy in 674 patients. Cancer. 2007;109(6):1068-77.

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

43. Berghmans T, Dusart M, Paesmans M, Hossein-Foucher C, Buvat I, Castaigne C, et al. Primary tumor standardized uptake value (SUVmax) measured on fluorodeoxyglucose positron emission tomography (FDG-PET) is of prognostic value for survival in non-small cell lung cancer (NSCLC): a systematic review and meta-analysis (MA) by the European Lung Cancer Working Party for the IASLC Lung Cancer Staging Project. J Thorac Oncol. 2008;3(1):6-12.

44. Downey RJ, Akhurst T, Gonen M, Park B, Rusch V. Fluorine-18 fluorodeoxyglucose positron emission tomographic maximal standardized uptake value predicts survival independent of clinical but not pathologic TNM staging of resected non-small cell lung cancer. Journal of Thoracic & Cardiovascular Surgery. 2007;133(6):1419-27.

45. Hoang JK, Hoagland LF, Coleman RE, Coan AD, Herndon JE, 2nd, Patz EF, Jr. Prognostic value of fluorine-18 fluorodeoxyglucose positron emission tomography imaging in patients with advanced-stage non-small-cell lung carcinoma. Journal of Clinical Oncology. 2008;26(9):1459-64.

46. Tanvetyanon T, Eikman EA, Sommers E, Robinson L, Boulware D, Bepler G. Computed tomography response, but not positron emission tomography scan response, predicts survival after neoadjuvant chemotherapy for resectable non-small-cell lung cancer. Journal of Clinical Oncology. 2008;26(28):4610-6.

47. Shin KM, Lee KS, Shim YM, Kim J, Kim BT, Kwon OJ, et al. FDG PET/CT and Mediastinal Nodal Metastasis Detection in Stage T1 Non-Small Cell Lung Cancer: Prognostic Implications. Korean Journal of Radiology. 2008;9(6):481-9.

48. Nguyen XC, Lee WW, Chung J-H, Park SY, Sung SW, Kim YK, et al. FDG uptake, glucose transporter type 1, and Ki-67 expressions in non-small-cell lung cancer: correlations and prognostic values. European Journal of Radiology. 2007;62(2):214-9.

49. Goodgame B, Pillot GA, Yang Z, Shriki J, Meyers BF, Zoole J, et al. Prognostic value of preoperative positron emission tomography in resected stage I non-small cell lung cancer. Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer. 2008;3(2):130-4.

50. Mac Manus MP, Hicks RJ, Matthews JP, Wirth A, Rischin D, Ball DL. Metabolic (FDG-PET) response after radical radiotherapy/chemoradiotherapy for non-small cell lung cancer correlates with patterns of failure. Lung Cancer. 2005;49(1):95-108.

51. Hoekstra CJ, Stroobants SG, Smit EF, Vansteenkiste J, van Tinteren H, Postmus PE, et al. Prognostic relevance of response evaluation using [18F]-2-fluoro-2-deoxy-D-glucose positron emission tomography in patients with locally advanced non-small-cell lung cancer. Journal of Clinical Oncology. 2005;23(33):8362-70.

52. Ohtsuka T, Nomori H, Watanabe K-I, Kaji M, Naruke T, Suemasu K, et al. Prognostic significance of [(18)F]fluorodeoxyglucose uptake on positron emission tomography in patients with pathologic stage I lung adenocarcinoma. Cancer. 2006;107(10):2468-73.

53. Hellwig D, Groschel A, Graeter TP, Hellwig AP, Nestle U, Schafers H-J, et al. Diagnostic performance and prognostic impact of FDG-PET in suspected recurrence of surgically treated non-small cell lung cancer. European Journal of Nuclear Medicine & Molecular Imaging. 2006;33(1):13-21.

54. Eschmann SM, Friedel G, Paulsen F, Reimold M, Hehr T, Budach W, et al. 18F-FDG PET for assessment of therapy response and preoperative re-evaluation after neoadjuvant radio-chemotherapy in stage III non-small cell lung cancer. European Journal of Nuclear Medicine & Molecular Imaging. 2007;34(4):463-71.

55. De Leyn P, Stroobants S, De Wever W, Lerut T, Coosemans W, Decker G, et al. Prospective comparative study of integrated positron emission tomography-computed tomography scan

compared with remediastinoscopy in the assessment of residual mediastinal lymph node disease after induction chemotherapy for mediastinoscopy-proven stage IIIA-N2 Non-small-cell lung cancer: a Leuven Lung Cancer Group Study. J Clin Oncol. 2006;24(21):3333-9.

56. Agency for Healthcare Research and Quality. Positron Emission Tomography for Nine Cancers (Bladder, Brain, Cervical, Kidney, Ovarian, Pancreatic, Prostate, Small Cell Lung, Testicular). Rockville: AHRQ; 2008.
57. Flores RM, Akhurst T, Gonen M, Zakowski M, Dycoco J, Larson SM, et al. Positron emission tomography predicts survival in malignant pleural mesothelioma. Journal of Thoracic & Cardiovascular Surgery. 2006;132(4):763-8.
58. National Institute for Health and Clinical Excellence. Improving Outcomes in Haematological Cancers. National Institute for Health and Clinical Excellence; 2003.
59. Pakos EE, Fotopoulos AD, Ioannidis JPA. 18F-FDG PET for evaluation of bone marrow infiltration in staging of lymphoma: a meta-analysis. Journal of Nuclear Medicine. 2005;46(6):958-63.
60. Tsujikawa T, Otsuka H, Morita N, Saegusa H, Kobayashi M, Okazawa H, et al. Does partial volume corrected maximum SUV based on count recovery coefficient in 3D-PET/CT correlate with clinical aggressiveness of non-Hodgkin's lymphoma? Annals of Nuclear Medicine. 2008;22(1):23-30.
61. Pelosi E, Penna D, Deandreis D, Chiappella A, Skanjeti A, Vitolo U, et al. FDG-PET in the detection of bone marrow disease in Hodgkin's disease and aggressive non-Hodgkin's lymphoma and its impact on clinical management. The Quarterly Journal of Nuclear Medicine & Molecular Imaging. 2008;52(1):9-16.
62. Bucerius J, Herkel C, Joe AY, Altehoefer C, Finke J, Moser E, et al. (18)F-FDG PET and conventional imaging for assessment of Hodgkin's disease and non Hodgkin's lymphoma. An analysis of 193 patient studies. Nuclear-Medizin. 2006;45(3):105-10; quiz N25-6.
63. Kwee TC, Kwee RM, Nievelstein RAJ. Imaging in staging of malignant lymphoma: a systematic review. Blood. 2008;111(2):504-16.
64. Terasawa T, Nihashi T, Hotta T, Nagai H. 18F-FDG PET for posttherapy assessment of Hodgkin's disease and aggressive Non-Hodgkin's lymphoma: a systematic review. Journal of Nuclear Medicine. 2008;49(1):13-21.
65. Zijlstra JM, Lindauer-van der Werf G, Hoekstra OS, Hooft L, Riphagen II, Huijgens PC. 18F-fluoro-deoxyglucose positron emission tomography for post-treatment evaluation of malignant lymphoma: a systematic review. Haematologica. 2006;91(4):522-9.
66. Janikova A, Bolcak K, Pavlik T, Mayer J, Kral Z. Value of [18F]fluorodeoxyglucose positron emission tomography in the management of follicular lymphoma: the end of a dilemma? Clin Lymphoma Myeloma. 2008;8(5):287-93.
67. Gallamini A, Hutchings M, Rigacci L, Specht L, Merli F, Hansen M, et al. Early interim 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. J Clin Oncol. 2007;25(24):3746-52.
68. Schot BW, Zijlstra JM, Sluiter WJ, van Imhoff GW, Pruim J, Vaalburg W, et al. Early FDG-PET assessment in combination with clinical risk scores determines prognosis in recurring lymphoma. Blood. 2007;109(2):486-91.
69. Gallamini A, Rigacci L, Merli F, Nassi L, Bosi A, Capodanno I, et al. The predictive value of positron emission tomography scanning performed after two courses of standard therapy on treatment outcome in advanced stage Hodgkin's disease. Haematologica. 2006;91(4):475-81.
70. Hutchings M, Loft A, Hansen M, Pedersen LM, Buhl T, Jurlander J, et al. FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. Blood. 2006;107(1):52-9.
71. Medical Services Advisory Committee. Positron emission tomography for head and neck cancer. Canberra: 2008.
72. Isles MG, McConkey C, Mehanna HM. A systematic review and meta-analysis of the role of positron emission tomography in the follow up of head and neck squamous cell carcinoma following radiotherapy or chemoradiotherapy. Clinical Otolaryngology. 2008;33(3):210-22.

73. Kyzas PA, Evangelou E, Denaxa-Kyza D, Ioannidis JPA. 18F-fluorodeoxyglucose positron emission tomography to evaluate cervical node metastases in patients with head and neck squamous cell carcinoma: a meta-analysis. J Natl Cancer Inst. 2008;100(10):712-20.
74. Brouwer J, Hooft L, Hoekstra OS, Riphagen II, Castelijns JA, de Bree R, et al. Systematic review: accuracy of imaging tests in the diagnosis of recurrent laryngeal carcinoma after radiotherapy. Head & Neck. 2008;30(7):889-97.
75. Liu T, Xu W, Yan W-L, Ye M, Bai Y-R, Huang G. FDG-PET, CT, MRI for diagnosis of local residual or recurrent nasopharyngeal carcinoma, which one is the best? A systematic review. Radiotherapy & Oncology. 2007;85(3):327-35.
76. Dong M-j, Zhao K, Lin X-t, Zhao J, Ruan L-x, Liu Z-f. Role of fluorodeoxyglucose-PET versus fluorodeoxyglucose-PET/computed tomography in detection of unknown primary tumor: a meta-analysis of the literature. Nuclear Medicine Communications. 2008;29(9):791-802.
77. Garin E, Prigent-Lejeune F, Lesimple T, Barge M-L, Rousseau C, Devillers A, et al. Impact of PET-FDG in the diagnosis and therapeutic care of patients presenting with metastases of unknown primary. Cancer Invest. 2007;25(4):232-9.
78. Johansen J, Buus S, Loft A, Keiding S, Overgaard M, Hansen HS, et al. Prospective study of 18FDG-PET in the detection and management of patients with lymph node metastases to the neck from an unknown primary tumor. Results from the DAHANCA-13 study. Head & Neck. 2008;30(4):471-8.
79. Agency for Health Technology Assessment in P. Cost-effectiveness analysis of PET-CT positron emission tomography and the diagnostic technologies financed from public sources in oncological diagnostics in Poland. Warsaw: Agency for Health Technology Assessment in Poland (AHTAPol). 2007:328.
80. Krabbe CA, Dijkstra PU, Pruim J, van der Laan BFM, van der Wal JE, Gravendeel JP, et al. FDG PET in oral and oropharyngeal cancer. Value for confirmation of N0 neck and detection of occult metastases. Oral Oncology. 2008;44(1):31-6.
81. Ng S-H, Chan S-C, Liao C-T, Chang JT-C, Ko S-F, Wang H-M, et al. Distant metastases and synchronous second primary tumors in patients with newly diagnosed oropharyngeal and hypopharyngeal carcinomas: evaluation of (18)F-FDG PET and extended-field multi-detector row CT. Neuroradiology. 2008;50(11):969-79.
82. Senft A, de Bree R, Hoekstra OS, Kuik DJ, Golding RP, Oyen WJG, et al. Screening for distant metastases in head and neck cancer patients by chest CT or whole body FDG-PET: a prospective multicenter trial. Radiotherapy & Oncology. 2008;87(2):221-9.
83. Horiuchi C, Taguchi T, Yoshida T, Nishimura G, Kawakami M, Tanigaki Y, et al. Early assessment of clinical response to concurrent chemoradiotherapy in head and neck carcinoma using fluoro-2-deoxy-d-glucose positron emission tomography. Auris, Nasus, Larynx. 2008;35(1):103-8.
84. Liu T, Xu W, Yan WL, Ye M, Bai YR, Huang G. FDG-PET, CT, MRI for diagnosis of local residual or recurrent nasopharyngeal carcinoma, which one is the best: a systematic review. Radiotherapy and Oncology. 2007;85(3):327-35.
85. Cicone F, Loose D, Deron P, Vermeersch H, Signore A, Van de Vyvere F, et al. Prognostic value of FDG uptake by the bone marrow in squamous cell carcinoma of the head and neck. Nuclear Medicine Communications. 2008;29(5):431-5.
86. Peeters M, Zlotta A, Roucoux F, De Greve J, Van Belle S, Haelterman M, et al. National Clinical Practice Guidelines of the College of Oncology : clinical practice guideline for colorectal cancer. Good Clinical Practice (GCP). Brussels: Belgian Health Care Knowledge Centre (KCE); 2006. KCE reports 29Suppl.
87. Penninckx F, Roels S, Leonard D, Laurent S, Decaestecker J, De Vleeschouwer C, et al. Quality of care for rectal cancer – Phase 1: a clinical practice guideline for rectal cancer. Good Clinical Practice (GCP). Brussels: Belgian Health Care Knowledge Centre (KCE); 2007. KCE reports 69.
88. Bipat S, van Leeuwen MS, Comans EFI, Pijl MEJ, Bossuyt PMM, Zwinderman AH, et al. Colorectal liver metastases: CT, MR imaging, and PET for diagnosis--meta-analysis. Radiology. 2005;237(1):123-31.

89. Tateishi U, Maeda T, Morimoto T, Miyake M, Arai Y, Kim EE. Non-enhanced CT versus contrast-enhanced CT in integrated PET/CT studies for nodal staging of rectal cancer. European Journal of Nuclear Medicine & Molecular Imaging. 2007;34(10):1627-34.

90. Llamas-Elvira JM, Rodriguez-Fernandez A, Gutierrez-Sainz J, Gomez-Rio M, Bellon-Guardia M, Ramos-Font C, et al. Fluorine-18 fluorodeoxyglucose PET in the preoperative staging of colorectal cancer. European Journal of Nuclear Medicine & Molecular Imaging. 2007;34(6):859-67.

91. Coenegrachts K, De Geeter F, ter Beek L, Walgraeve N, Bipat S, Stoker J, et al. Comparison of MRI (including SS SE-EPI and SPIO-enhanced MRI) and FDG-PET/CT for the detection of colorectal liver metastases. European Radiology. 2009;19(2):370-9.

92. Kong G, Jackson C, Koh DM, Lewington V, Sharma B, Brown G, et al. The use of 18F-FDG PET/CT in colorectal liver metastases--comparison with CT and liver MRI. European Journal of Nuclear Medicine & Molecular Imaging. 2008;35(7):1323-9.

93. Chua SC, Groves AM, Kayani I, Menezes L, Gacinovic S, Du Y, et al. The impact of 18F-FDG PET/CT in patients with liver metastases. European Journal of Nuclear Medicine & Molecular Imaging. 2007;34(12):1906-14.

94. Cascini GL, Avallone A, Delrio P, Guida C, Tatangelo F, Marone P, et al. 18F-FDG PET is an early predictor of pathologic tumor response to preoperative radiochemotherapy in locally advanced rectal cancer. Journal of Nuclear Medicine. 2006;47(8):1241-8.

95. Rosenberg R, Herrmann K, Gertler R, Kunzli B, Essler M, Lordick F, et al. The predictive value of metabolic response to preoperative radiochemotherapy in locally advanced rectal cancer measured by PET/CT. International Journal of Colorectal Disease. 2009;24(2):191-200.

96. Capirci C, Rampin L, Erba PA, Galeotti F, Crepaldi G, Banti E, et al. Sequential FDG-PET/CT reliably predicts response of locally advanced rectal cancer to neo-adjuvant chemo-radiation therapy. European Journal of Nuclear Medicine & Molecular Imaging. 2007;34(10):1583-93.

97. Zhang C, Chen Y, Xue H, Zheng P, Tong J, Liu J, et al. Diagnostic value of FDG-PET in recurrent colorectal carcinoma: a meta-analysis. International Journal of Cancer. 2009;124(1):167-73.

98. Wiering B, Krabbe PFM, Jager GJ, Oyen WJG, Ruers TJM. The impact of fluor-18-deoxyglucose-positron emission tomography in the management of colorectal liver metastases. Cancer. 2005;104(12):2658-70.

99. Sobhani I, Tiret E, Lebtahi R, Aparicio T, Itti E, Montravers F, et al. Early detection of recurrence by 18FDG-PET in the follow-up of patients with colorectal cancer. Br J Cancer. 2008;98(5):875-80.

100. Votrubova J, Belohlavek O, Jaruskova M, Oliverius M, Lohynska R, Trskova K, et al. The role of FDG-PET/CT in the detection of recurrent colorectal cancer. European Journal of Nuclear Medicine & Molecular Imaging. 2006;33(7):779-84.

101. Bellomi M, Rizzo S, Travaini LL, Bazzi L, Trifiro G, Zampino MG, et al. Role of multidetector CT and FDG-PET/CT in the diagnosis of local and distant recurrence of resected rectal cancer. Radiologia Medica. 2007;112(5):681-90.

102. de Geus-Oei LF, van Laarhoven HWM, Visser EP, Hermsen R, van Hoorn BA, Kamm YJL, et al. Chemotherapy response evaluation with FDG-PET in patients with colorectal cancer. Annals of Oncology. 2008;19(2):348-52.

103. Nakagawa K, Yamashita H, Nakamura N, Igaki H, Tago M, Hosoi Y, et al. Preoperative radiation response evaluated by 18-fluorodeoxyglucose positron emission tomography predicts survival in locally advanced rectal cancer. Dis Colon Rectum. 2008;51(7):1055-60.

104. Kalff V, Duong C, Drummond EG, Matthews JP, Hicks RJ. Findings on 18F-FDG PET scans after neoadjuvant chemoradiation provides prognostic stratification in patients with locally advanced rectal carcinoma subsequently treated by radical surgery. Journal of Nuclear Medicine. 2006;47(1):14-22.

105. College of Oncology. National guidelines melanoma. Available from: https://portal.health.fgov.be/pls/portal/docs/PAGE/INTERNET_PG/HOMEPAGE_MENU/GEZONDHEIDZORG1_MENU/OVERLEGSTRUCTUREN1_MENU/COLLEGESVANGENEESHEREN1_MENU/ONCOLOGIE1_MENU/RICHTLIJNENKLINISCHEPRAKTIJK1_HIDE/RICHTLIJNENKLINISCHEPRAKTIJK1_DOCS/MELANOMA%20V1.2007%20(NL)_0.PDF.

106. Krug B, Crott R, Lonneux M, Baurain J-F, Pirson A-S, Vander Borght T. Role of PET in the initial staging of cutaneous malignant melanoma: systematic review. Radiology. 2008;249(3):836-44.

107. Singh B, Ezziddin S, Palmedo H, Reinhardt M, Strunk H, Tuting T, et al. Preoperative 18F-FDG-PET/CT imaging and sentinel node biopsy in the detection of regional lymph node metastases in malignant melanoma. Melanoma Research. 2008;18(5):346-52.

108. Strobel K, Skalsky J, Kalff V, Baumann K, Seifert B, Joller-Jemelka H, et al. Tumour assessment in advanced melanoma: value of FDG-PET/CT in patients with elevated serum S-100B. European Journal of Nuclear Medicine & Molecular Imaging. 2007;34(9):1366-75.

109. Akcali C, Zincirkeser S, Erbagcy Z, Akcali A, Halac M, Durak G, et al. Detection of metastases in patients with cutaneous melanoma using FDG-PET/CT. Journal of International Medical Research. 2007;35(4):547-53.

110. Strobel K, Dummer R, Husarik DB, Perez Lago M, Hany TF, Steinert HC. High-risk melanoma: accuracy of FDG PET/CT with added CT morphologic information for detection of metastases. Radiology. 2007;244(2):566-74.

111. Medical Services Advisory Committee. Positron emission tomography for recurrent melanoma. Canberra: MSAC; 2008.

112. Imbriaco M, Caprio MG, Limite G, Pace L, De Falco T, Capuano E, et al. Dual-time-point 18F-FDG PET/CT versus dynamic breast MRI of suspicious breast lesions. AJR. 2008;American Journal of Roentgenology. 191(5):1323-30.

113. Agency for Healthcare Research and Quality. Effectiveness of Noninvasive Diagnostic Tests for Breast Abnormalities. Rockville: AHRQ; 2006.

114. Bourguet P, Hitzel A, Houvenaeghel G, Vinatier D, Bosquet L. 2005 monitoring report: use of positron emission tomography with fluorodeoxyglucose in the management of patients with breast cancer, ovarian cancer, and uterine cancer. Gynecol Obstet Fertil. 2006;34(5):437-59.

115. Shie P, Cardarelli R, Brandon D, Erdman W, Abdulrahim N. Meta-analysis: comparison of F-18 Fluorodeoxyglucose-positron emission tomography and bone scintigraphy in the detection of bone metastases in patients with breast cancer.[erratum appears in Clin Nucl Med. 2008 May;33(5):329]. Clinical Nuclear Medicine. 2008;33(2):97-101.

116. Sloka JS, Hollett PD, Mathews M. A quantitative review of the use of FDG-PET in the axillary staging of breast cancer. Medical Science Monitor. 2007;13(3):RA37-46.

117. Agency for Healthcare Research and Quality. FDG positron emission tomography for evaluating breast cancer - systematic review. Rockville: Agency for Healthcare Research and Quality; 2001.

118. Gil-Rendo A, Zornoza G, Garcia-Velloso MJ, Regueira FM, Beorlegui C, Cervera M. Fluorodeoxyglucose positron emission tomography with sentinel lymph node biopsy for evaluation of axillary involvement in breast cancer. British Journal of Surgery. 2006;93(6):707-12.

119. Kumar R, Zhuang H, Schnall M, Conant E, Damia S, Weinstein S, et al. FDG PET positive lymph nodes are highly predictive of metastasis in breast cancer. Nuclear Medicine Communications. 2006;27(3):231-6.

120. Ueda S, Tsuda H, Asakawa H, Omata J, Fukatsu K, Kondo N, et al. Utility of 18F-fluoro-deoxyglucose emission tomography/computed tomography fusion imaging (18F-FDG PET/CT) in combination with ultrasonography for axillary staging in primary breast cancer. BMC Cancer. 2008;8:165.

121. Veronesi U, De Cicco C, Galimberti VE, Fernandez JR, Rotmensz N, Viale G, et al. A comparative study on the value of FDG-PET and sentinel node biopsy to identify occult axillary metastases. Annals of Oncology. 2007;18(3):473-8.

122. Shie P, Cardarelli R, Brandon D, Erdman W, Rahim NA. Meta-analysis: comparison of F-18 fluorodeoxyglucose-positron emission tomography and bone scintigraphy in the detection of bone metastases in patients with breast cancer. Clinical Nuclear Medicine. 2008;33(2):97-101.

123. Nakai T, Okuyama C, Kubota T, Yamada K, Ushijima Y, Taniike K, et al. Pitfalls of FDG-PET for the diagnosis of osteoblastic bone metastases in patients with breast cancer.[see comment]. European Journal of Nuclear Medicine & Molecular Imaging. 2005;32(11):1253-8.

124. Schwarz-Dose J, Untch M, Tiling R, Sassen S, Mahner S, Kahlert S, et al. Monitoring primary systemic therapy of large and locally advanced breast cancer by using sequential positron emission tomography imaging with [18F]fluorodeoxyglucose. Journal of Clinical Oncology. 2009;27(4):535-41.

125. Berriolo-Riedinger A, Touzery C, Riedinger J-M, Toubeau M, Coudert B, Arnould L, et al. [18F]FDG-PET predicts complete pathological response of breast cancer to neoadjuvant chemotherapy. European Journal of Nuclear Medicine & Molecular Imaging. 2007;34(12):1915-24.

126. Cachin F, Prince HM, Hogg A, Ware RE, Hicks RJ. Powerful prognostic stratification by [18F]fluorodeoxyglucose positron emission tomography in patients with metastatic breast cancer treated with high-dose chemotherapy. Journal of Clinical Oncology. 2006;24(19):3026-31.

127. Emmering J, Krak NC, Van der Hoeven JJM, Spreeuwenberg MD, Twisk JWR, Meijer S, et al. Preoperative [18F] FDG-PET after chemotherapy in locally advanced breast cancer: prognostic value as compared with histopathology. Annals of Oncology. 2008;19(9):1573-7.

128. Peeters M, Lerut T, Vlayen J, Mambourg F, Ectors N, Deprez P, et al. [Guideline for the esophageal and gastric cancer: scientific support of the College of Onclogy]. Good Clinical Practice (GCP). Brussels: Belgian Health Care Knowledge Centre (KCE); 2008. KCE Reports 75 .

129. Medical Services Advisory Committee. Positron emission tomography for oesophageal and gastric cancer. Canberra: MSAC; 2008.

130. van Vliet EPM, Heijenbrok-Kal MH, Hunink MGM, Kuipers EJ, Siersema PD. Staging investigations for oesophageal cancer: a meta-analysis. British Journal of Cancer. 2008;98(3):547-57.

131. Westerterp M, Van Westreenen HL, Sloof GW, Plukker JTM, Van Lanschot JJB. Role of positron emission tomography in the (re-)staging of oesophageal cancer. Scandinavian Journal of Gastroenterology - Supplement. 2006(243):116-22.

132. van Westreenen HL, Westerterp M, Bossuyt PMM, Pruim J, Sloof GW, van Lanschot JJB, et al. Systematic review of the staging performance of 18F-fluorodeoxyglucose positron emission tomography in esophageal cancer. J Clin Oncol. 2004;22(18):3805-12.

133. Huguier M, Barrier A, Zacharias T, Valinas R. [Positron emission tomography of gastrointestinal carcinomas]. Bulletin de l Academie Nationale de Medecine. 2006;190(1):75-84.

134. Kato H, Miyazaki T, Nakajima M, Takita J, Kimura H, Faried A, et al. Comparison between whole-body positron emission tomography and bone scintigraphy in evaluating bony metastases of esophageal carcinomas. Anticancer Research. 2005;25(6C):4439-44.

135. Little SG, Rice TW, Bybel B, Mason DP, Murthy SC, Falk GW, et al. Is FDG-PET indicated for superficial esophageal cancer? European Journal of Cardio-Thoracic Surgery. 2007;31(5):791-6.

136. Schreurs LMA, Verhoef CCPM, van der Jagt EJ, van Dam GM, Groen H, Plukker JTM. Current relevance of cervical ultrasonography in staging cancer of the esophagus and gastroesophageal junction. European Journal of Radiology. 2008;67(1):105-11.

137. Higuchi I, Yasuda T, Yano M, Doki Y, Miyata H, Tatsumi M, et al. Lack of fludeoxyglucose F 18 uptake in posttreatment positron emission tomography as a significant predictor of survival after subsequent surgery in multimodality treatment for patients with locally advanced esophageal squamous cell carcinoma. Journal of Thoracic & Cardiovascular Surgery. 2008;136(1):205-12.

138. Song SY, Kim JH, Ryu JS, Lee GH, Kim SB, Park SI, et al. FDG-PET in the prediction of pathologic response after neoadjuvant chemoradiotherapy in locally advanced, resectable esophageal cancer. International Journal of Radiation Oncology, Biology, Physics. 2005;63(4):1053-9.

139. Wieder HA, Ott K, Lordick F, Becker K, Stahl A, Herrmann K, et al. Prediction of tumor response by FDG-PET: comparison of the accuracy of single and sequential studies in patients with adenocarcinomas of the esophagogastric junction. European Journal of Nuclear Medicine & Molecular Imaging. 2007;34(12):1925-32.

140. Roedl JB, Colen RR, Holalkere NS, Fischman AJ, Choi NC, Blake MA. Adenocarcinomas of the esophagus: response to chemoradiotherapy is associated with decrease of metabolic tumor

volume as measured on PET-CT. Comparison to histopathologic and clinical response evaluation. Radiotherapy & Oncology. 2008;89(3):278-86.

141. Erasmus JJ, Munden RF, Truong MT, Ho JJ, Hofstetter WL, Macapinlac HA, et al. Preoperative chemo-radiation-induced ulceration in patients with esophageal cancer: a confounding factor in tumor response assessment in integrated computed tomographic-positron emission tomographic imaging. Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer. 2006;1(5):478-86.
142. Lordick F, Ott K, Krause B-J, Weber WA, Becker K, Stein HJ, et al. PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial. Lancet Oncol. 2007;8(9):797-805.
143. Guo H, Zhu H, Xi Y, Zhang B, Li L, Huang Y, et al. Diagnostic and prognostic value of 18F-FDG PET/CT for patients with suspected recurrence from squamous cell carcinoma of the esophagus. Journal of Nuclear Medicine. 2007;48(8):1251-8.
144. Blackstock AW, Farmer MR, Lovato J, Mishra G, Melin SA, Oaks T, et al. A prospective evaluation of the impact of 18-F-fluoro-deoxy-D-glucose positron emission tomography staging on survival for patients with locally advanced esophageal cancer. Int J Radiat Oncol Biol Phys. 2006;64(2):455-60.
145. Cerfolio RJ, Bryant AS. Maximum standardized uptake values on positron emission tomography of esophageal cancer predicts stage, tumor biology, and survival. Ann Thorac Surg. 2006;82(2):391-4.
146. Cheze-Le Rest C, Metges J-P, Teyton P, Jestin-Le Tallec V, Lozac'h P, Volant A, et al. Prognostic value of initial fluorodeoxyglucose-PET in esophageal cancer: a prospective study. Nuclear Medicine Communications. 2008;29(7):628-35.
147. Choi JY, Jang K-T, Shim YM, Kim K, Ahn G, Lee K-H, et al. Prognostic significance of vascular endothelial growth factor expression and microvessel density in esophageal squamous cell carcinoma: comparison with positron emission tomography. Ann Surg Oncol. 2006;13(8):1054-62.
148. Chung HW, Lee K-H, Lee EJ, Lee SJ, Cho YS, Choi JY, et al. Comparison of uptake characteristics and prognostic value of 201Tl and 18F-FDG in esophageal cancer. World J Surg. 2008;32(1):69-75.
149. Duong CP, Hicks RJ, Weih L, Drummond E, Leong T, Michael M, et al. FDG-PET status following chemoradiotherapy provides high management impact and powerful prognostic stratification in oesophageal cancer. European Journal of Nuclear Medicine & Molecular Imaging. 2006;33(7):770-8.
150. Kim MK, Ryu J-S, Kim S-B, Ahn JH, Kim SY, Park S-I, et al. Value of complete metabolic response by (18)F-fluorodeoxyglucose-positron emission tomography in oesophageal cancer for prediction of pathologic response and survival after preoperative chemoradiotherapy. Eur J Cancer. 2007;43(9):1385-91.
151. Makino T, Doki Y, Miyata H, Yasuda T, Yamasaki M, Fujiwara Y, et al. Use of (18)F-fluorodeoxyglucose-positron emission tomography to evaluate responses to neo-adjuvant chemotherapy for primary tumor and lymph node metastasis in esophageal squamous cell carcinoma. Surgery. 2008;144(5):793-802.
152. Omloo JMT, Sloof GW, Boellaard R, Hoekstra OS, Jager PL, van Dullemen HM, et al. Importance of fluorodeoxyglucose-positron emission tomography (FDG-PET) and endoscopic ultrasonography parameters in predicting survival following surgery for esophageal cancer. Endoscopy. 2008;40(6):464-71.
153. Ott K, Weber WA, Lordick F, Becker K, Busch R, Herrmann K, et al. Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction. Journal of Clinical Oncology. 2006;24(29):4692-8.
154. Roedl JB, Halpern EF, Colen RR, Sahani DV, Fischman AJ, Blake MA. Metabolic Tumor Width Parameters as Determined on PET/CT Predict Disease-free Survival and Treatment Response in Squamous Cell Carcinoma of the Esophagus. Molecular Imaging & Biology. 2009;11(1):54-60.
155. Sun L, Su X-H, Guan Y-S, Pan W-M, Luo Z-M, Wei J-H, et al. Clinical role of 18F-fluorodeoxyglucose positron emission tomography/computed tomography in post-operative follow up of gastric cancer: initial results. World Journal of Gastroenterology. 2008;14(29):4627-32.

156. Ott K, Herrmann K, Lordick F, Wieder H, Weber WA, Becker K, et al. Early metabolic response evaluation by fluorine-18 fluorodeoxyglucose positron emission tomography allows in vivo testing of chemosensitivity in gastric cancer: long-term results of a prospective study. Clinical Cancer Research. 2008;14(7):2012-8.

157. Di Fabio F, Pinto C, Rojas Llimpe FL, Fanti S, Castellucci P, Longobardi C, et al. The predictive value of 18F-FDG-PET early evaluation in patients with metastatic gastric adenocarcinoma treated with chemotherapy plus cetuximab. Gastric Cancer. 2007;10(4):221-7.

158. Association of Comprehensive Cancer Centres. Thyroid carcinoma. Nation-wide guideline. Version 1.0. ACCC; 2007.

159. Sebastianes FM, Cerci JJ, Zanoni PH, Soares J, Jr., Chibana LK, Tomimori EK, et al. Role of 18F-fluorodeoxyglucose positron emission tomography in preoperative assessment of cytologically indeterminate thyroid nodules. Journal of Clinical Endocrinology & Metabolism. 2007;92(11):4485-8.

160. de Geus-Oei L-F, Pieters GFFM, Bonenkamp JJ, Mudde AH, Bleeker-Rovers CP, Corstens FHM, et al. 18F-FDG PET reduces unnecessary hemithyroidectomies for thyroid nodules with inconclusive cytologic results. Journal of Nuclear Medicine. 2006;47(5):770-5.

161. Kresnik E, Gallowitsch HJ, Mikosch P, Stettner H, Igerc I, Gomez I, et al. Fluorine-18-fluorodeoxyglucose positron emission tomography in the preoperative assessment of thyroid nodules in an endemic goiter area. Surgery. 2003;133(3):294-9.

162. Alzahrani AS, Mohamed GE, Al Rifai A, Al-Sugair A, Abdel Salam SA, Sulaiman OM, et al. Role of [18F]fluorodeoxyglucose positron emission tomography in follow-up of differentiated thyroid cancer. Endocrine Practice. 2006;12(2):152-8.

163. Choi M-Y, Chung J-K, Lee H-Y, So Y, Park DJ, Jeong JM, et al. The clinical impact of 18F-FDG PET in papillary thyroid carcinoma with a negative 1311 whole body scan: a single-center study of 108 patients. Annals of Nuclear Medicine. 2006;20(8):547-52.

164. Iagaru A, Masamed R, Singer PA, Conti PS. 2-Deoxy-2-[18F]fluoro-D-glucose-positron emission tomography and positron emission tomography/computed tomography diagnosis of patients with recurrent papillary thyroid cancer. Molecular Imaging & Biology. 2006;8(5):309-14.

165. Pryma DA, Schoder H, Gonen M, Robbins RJ, Larson SM, Yeung HWD. Diagnostic accuracy and prognostic value of 18F-FDG PET in Hurthle cell thyroid cancer patients. Journal of Nuclear Medicine. 2006;47(8):1260-6.

166. Finkelstein SE, Grigsby PW, Siegel BA, Dehdashti F, Moley JF, Hall BL. Combined [18F]Fluorodeoxyglucose positron emission tomography and computed tomography (FDG-PET/CT) for detection of recurrent, 131I-negative thyroid cancer. Annals of Surgical Oncology. 2008;15(1):286-92.

167. Shammas A, Degirmenci B, Mountz JM, McCook BM, Branstetter B, Bencherif B, et al. 18F-FDG PET/CT in patients with suspected recurrent or metastatic well-differentiated thyroid cancer. Journal of Nuclear Medicine. 2007;48(2):221-6.

168. Freudenberg LS, Antoch G, Frilling A, Jentzen W, Rosenbaum SJ, Kuhl H, et al. Combined metabolic and morphologic imaging in thyroid carcinoma patients with elevated serum thyroglobulin and negative cervical ultrasonography: role of 124I-PET/CT and FDG-PET. European Journal of Nuclear Medicine & Molecular Imaging. 2008;35(5):950-7.

169. Freudenberg LS, Frilling A, Kuhl H, Muller SP, Jentzen W, Bockisch A, et al. Dual-modality FDG-PET/CT in follow-up of patients with recurrent iodine-negative differentiated thyroid cancer. European Radiology. 2007;17(12):3139-47.

170. Robbins RJ, Wan Q, Grewal RK, Reibke R, Gonen M, Strauss HW, et al. Real-time prognosis for metastatic thyroid carcinoma based on 2-[18F]fluoro-2-deoxy-D-glucose-positron emission tomography scanning. Journal of Clinical Endocrinology & Metabolism. 2006;91(2):498-505.

171. Peeters M, Vlayen J, Stordeur S, Mambourg F, Boterberg T, de Hemptinne B, et al. Wetenschappelijke ondersteuning van het College voor Oncologie: een nationale praktijkrichtlijn voor de aanpak van pancreaskanker. Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2009. KCE reports 105A.

172. Singer E, Gschwantler M, Plattner D, Kriwanek S, Armbruster C, Schueller J, et al. Differential diagnosis of benign and malign pancreatic masses with 18F-fluordeoxyglucose-positron

emission tomography recorded with a dual-head coincidence gamma camera. European Journal of Gastroenterology & Hepatology. 2007;19(6):471-8.

173. Sperti C, Pasquali C, Decet G, Chierichetti F, Liessi G, Pedrazzoli S. F-18-fluorodeoxyglucose positron emission tomography in differentiating malignant from benign pancreatic cysts: a prospective study. Journal of Gastrointestinal Surgery. 2005;9(1):22-8.

174. Schick V, Franzius C, Beyna T, Oei ML, Schnekenburger J, Weckesser M, et al. Diagnostic impact of 18F-FDG PET-CT evaluating solid pancreatic lesions versus endosonography, endoscopic retrograde cholangio-pancreatography with intraductal ultrasonography and abdominal ultrasound. European Journal of Nuclear Medicine & Molecular Imaging. 2008;35(10):1775-85.

175. Strobel K, Heinrich S, Bhure U, Soyka J, Veit-Haibach P, Pestalozzi BC, et al. Contrast-enhanced 18F-FDG PET/CT: 1-stop-shop imaging for assessing the resectability of pancreatic cancer. Journal of Nuclear Medicine. 2008;49(9):1408-13.

176. Nishiyama Y, Yamamoto Y, Monden T, Sasakawa Y, Tsutsui K, Wakabayashi H, et al. Evaluation of delayed additional FDG PET imaging in patients with pancreatic tumour. Nuclear Medicine Communications. 2005;26(10):895-901.

177. Lyshchik A, Higashi T, Nakamoto Y, Fujimoto K, Doi R, Imamura M, et al. Dual-phase 18F-fluoro-2-deoxy-D-glucose positron emission tomography as a prognostic parameter in patients with pancreatic cancer. Eur J Nucl Med Mol Imaging. 2005;32(4):389-97.

178. Yoon KT, Kim JK, Kim DY, Ahn SH, Lee JD, Yun M, et al. Role of 18F-fluorodeoxyglucose positron emission tomography in detecting extrahepatic metastasis in pretreatment staging of hepatocellular carcinoma. Oncology. 2007;72 Suppl 1:104-10.

179. Ho C-l, Chen S, Yeung DWC, Cheng TKC. Dual-tracer PET/CT imaging in evaluation of metastatic hepatocellular carcinoma. Journal of Nuclear Medicine. 2007;48(6):902-9.

180. Paudyal B, Paudyal P, Oriuchi N, Tsushima Y, Nakajima T, Endo K. Clinical implication of glucose transport and metabolism evaluated by 18F-FDG PET in hepatocellular carcinoma. Int J Oncol. 2008;33(5):1047-54.

181. Seo S, Hatano E, Higashi T, Hara T, Tada M, Tamaki N, et al. Fluorine-18 fluorodeoxyglucose positron emission tomography predicts tumor differentiation, P-glycoprotein expression, and outcome after resection in hepatocellular carcinoma. Clin Cancer Res. 2007;13(2 Pt 1):427-33.

182. Scottish Intercollegiate Guidelines Network. Management of cervical cancer. A national clinical guideline. Edinburgh: SIGN; 2008.

183. Selman TJ, Mann C, Zamora J, Appleyard T-L, Khan K. Diagnostic accuracy of tests for lymph node status in primary cervical cancer: a systematic review and meta-analysis. CMAJ. 2008;178(7):855-62.

184. Havrilesky L, Kulasingam SL, Matchar DB, Myers ER. FDG-PET for management of cervical and ovarian cancer. Gynecologic Oncology. 2005;97(1):183-91.

185. Kitajima K, Murakami K, Yamasaki E, Domeki Y, Kaji Y, Morita S, et al. Performance of integrated FDG-PET/contrast-enhanced CT in the diagnosis of recurrent uterine cancer: comparison with PET and enhanced CT. Eur J Nucl Med Mol Imaging. 2009;36(3):362-72.

186. Kitajima K, Murakami K, Yamasaki E, Domeki Y, Kaji Y, Sugimura K. Performance of FDG-PET/CT for diagnosis of recurrent uterine cervical cancer. European Radiology. 2008;18(10):2040-7.

187. van der Veldt AAM, Buist MR, van Baal MW, Comans EF, Hoekstra OS, Molthoff CFM. Clarifying the diagnosis of clinically suspected recurrence of cervical cancer: impact of 18F-FDG PET. Journal of Nuclear Medicine. 2008;49(12):1936-43.

188. Yen T-C, See L-C, Lai C-H, Tsai C-S, Chao A, Hsueh S, et al. Standardized uptake value in para-aortic lymph nodes is a significant prognostic factor in patients with primary advanced squamous cervical cancer. European Journal of Nuclear Medicine & Molecular Imaging. 2008;35(3):493-501.

189. Xue F, Lin LL, Dehdashti F, Miller TR, Siegel BA, Grigsby PW. F-18 fluorodeoxyglucose uptake in primary cervical cancer as an indicator of prognosis after radiation therapy. Gynecologic Oncology. 2006;101(1):147-51.

190. Kidd EA, Grigsby PW. Intratumoral metabolic heterogeneity of cervical cancer. Clin Cancer Res. 2008;14(16):5236-41.

191. Schwarz JK, Siegel BA, Dehdashti F, Grigsby PW. Association of posttherapy positron emission tomography with tumor response and survival in cervical carcinoma. JAMA. 2007;298(19):2289-95.

192. Agency for Healthcare Research and Quality. Management of Adnexal Mass. Rockville: AHRQ; 2006.

193. Yamamoto Y, Oguri H, Yamada R, Maeda N, Kohsaki S, Fukaya T. Preoperative evaluation of pelvic masses with combined 18F-fluorodeoxyglucose positron emission tomography and computed tomography. International Journal of Gynaecology & Obstetrics. 2008;102(2):124-7.

194. Kitajima K, Murakami K, Yamasaki E, Domeki Y, Kaji Y, Fukasawa I, et al. Performance of integrated FDG-PET/contrast-enhanced CT in the diagnosis of recurrent ovarian cancer: comparison with integrated FDG-PET/non-contrast-enhanced CT and enhanced CT. European Journal of Nuclear Medicine & Molecular Imaging. 2008;35(8):1439-48.

195. Torizuka T, Nakamura F, Takekuma M, Kanno T, Ogusu T, Yoshikawa E, et al. FDG PET for the assessment of myometrial infiltration in clinical stage I uterine corpus cancer. Nuclear Medicine Communications. 2006;27(6):481-7.

196. Kitajima K, Murakami K, Yamasaki E, Fukasawa I, Inaba N, Kaji Y, et al. Accuracy of 18F-FDG PET/CT in detecting pelvic and paraaortic lymph node metastasis in patients with endometrial cancer. AJR. 2008;American Journal of Roentgenology. 190(6):1652-8.

197. Chung HH, Kang WJ, Kim JW, Park N-H, Song Y-S, Chung J-K, et al. The clinical impact of [(18)F]FDG PET/CT for the management of recurrent endometrial cancer: correlation with clinical and histological findings. European Journal of Nuclear Medicine & Molecular Imaging. 2008;35(6):1081-8.

198. Corgna E, Betti M, Gatta G, et al. Renal Cancer. Crit Rev Oncol. 2007;64(3):247-62.

199. Bourguet P, Planchamp F, Montravers F, Vincendeau S. Utilisation de la TEP-FDG dans les cancers uro-néphrologiques - Bulletin de synthèse de veille 2006. FNCLCC; 2006.

200. Park JW, Jo MK, Lee HM. Significance of 18F-fluorodeoxyglucose positron-emission tomography/computed tomography for the postoperative surveillance of advanced renal cell carcinoma. BJU Int. 2009;103(5):615-9.

201. Bourguet P, Planchamp F, Montravers F, Vincendeau S, Association francaise du, Federation nationale des centres de lutte contre le c. 2006 technology monitoring report: clinical practice guideline: use of FDG-PET in kidney, prostate, testis and bladder cancers. Prog Urol. 2007;17(2):172-5.

202. de Wit M, Brenner W, Hartmann M, Kotzerke J, Hellwig D, Lehmann J, et al. [18F]-FDG-PET in clinical stage I/II non-seminomatous germ cell tumours: results of the German multicentre trial. Annals of Oncology. 2008;19(9):1619-23.

203. Oechsle K, Hartmann M, Brenner W, Venz S, Weissbach L, Franzius C, et al. [18F]Fluorodeoxyglucose positron emission tomography in nonseminomatous germ cell tumors after chemotherapy: the German multicenter positron emission tomography study group. Journal of Clinical Oncology. 2008;26(36):5930-5.

204. Popperl G, Kreth FW, Mehrkens JH, Herms J, Seelos K, Koch W, et al. FET PET for the evaluation of untreated gliomas: correlation of FET uptake and uptake kinetics with tumour grading. Eur J Nucl Med Mol Imaging. 2007;34(12):1933-42.

205. Roessler K, Gatterbauer B, Becherer A, Paul M, Kletter K, Prayer D, et al. Surgical target selection in cerebral glioma surgery: linking methionine (MET) PET image fusion and neuronavigation. Minim Invasive Neurosurg. 2007;50(5):273-80.

206. McCarthy M, Yuan JB, Campbell A, Lenzo NP, Butler-Henderson K. 18F-fluorodeoxyglucose positron emission tomography imaging in brain tumours: the Western Australia positron emission tomography/cyclotron service experience. Journal of Medical Imaging & Radiation Oncology. 2008;52(6):564-9.

207. Popperl G, Gotz C, Rachinger W, Schnell O, Gildehaus FJ, Tonn JC, et al. Serial O-(2-[(18)F]fluoroethyl)-L: -tyrosine PET for monitoring the effects of intracavitary radioimmunotherapy in patients with malignant glioma. Eur J Nucl Med Mol Imaging. 2006;33(7):792-800.

208. Rachinger W, Goetz C, Popperl G, Gildehaus FJ, Kreth FW, Holtmannspotter M, et al. Positron emission tomography with O-(2-[18F]fluoroethyl)-l-tyrosine versus magnetic resonance imaging in the diagnosis of recurrent gliomas. Neurosurgery. 2005;57(3):505-11.

209. Ceyssens S, Van Laere K, de Groot T, Goffin J, Bormans G, Mortelmans L. [11C]methionine PET, histopathology, and survival in primary brain tumors and recurrence. AJNR Am J Neuroradiol. 2006;27(7):1432-7.

210. Spence AM, Muzi M, Swanson KR, O'Sullivan F, Rockhill JK, Rajendran JG, et al. Regional hypoxia in glioblastoma multiforme quantified with [18F]fluoromisonidazole positron emission tomography before radiotherapy: correlation with time to progression and survival. Clin Cancer Res. 2008;14(9):2623-30.

211. Nandalur KR, Dwamena BA, Choudhri AF, Nandalur SR, Reddy P, Carlos RC. Diagnostic performance of positron emission tomography in the detection of coronary artery disease: a meta-analysis. Academic Radiology. 2008;15(4):444-51.

212. Beanlands RSB, Chow BJW, Dick A, Friedrich MG, Gulenchyn KY, Kiess M, et al. CCS/CAR/CANM/CNCS/CanSCMR joint position statement on advanced noninvasive cardiac imaging using positron emission tomography, magnetic resonance imaging and multidetector computed tomographic angiography in the diagnosis and evaluation of ischemic heart disease - Executive summary. Canadian Journal of Cardiology. 2007;23(2):107-19.

213. Medical Advisory S, Ontario Ministry of Health and Long-Term C. Positron emission tomography for the assessment of myocardial viability. Toronto: Medical Advisory Secretariat, Ontario Ministry of Health and Long-Term Care (MAS). 2005.

214. Beanlands RSB, Nichol G, Huszti E, Humen D, Racine N, Freeman M, et al. F-18-fluorodeoxyglucose positron emission tomography imaging-assisted management of patients with severe left ventricular dysfunction and suspected coronary disease: a randomized, controlled trial (PARR-2).[see comment]. J Am Coll Cardiol. 2007;50(20):2002-12.

215. Tio RA, Dabeshlim A, Siebelink H-MJ, de Sutter J, Hillege HL, Zeebregts CJ, et al. Comparison between the prognostic value of left ventricular function and myocardial perfusion reserve in patients with ischemic heart disease. J Nucl Med. 2009;50(2):214-9.

216. Santana CA, Faber TL, Soler-Peter M, Sanyal R, Esteves FP, Ornelas M, et al. Prognostic performance of quantitative PET tools for stratification of patients with ischemic cardiomyopathy undergoing myocardial viability assessment. Nucl Med Commun. 2008;29(11):970-81.

217. Suchowersky O, Reich S, Perlmutter J, Zesiewicz T, Gronseth G, Weiner WJ, et al. Practice Parameter: diagnosis and prognosis of new onset Parkinson disease (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology. 2006;66(7):968-75.

218. Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. Arch Neurol. 1999;56(1):33-9.

219. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry. 1992;55(3):181-4.

220. Heiss WD, Hilker R. The sensitivity of 18-fluorodopa positron emission tomography and magnetic resonance imaging in Parkinson's disease. Eur J Neurol. 2004;11(1):5-12.

221. Eckert T, Barnes A, Dhawan V, Frucht S, Gordon MF, Feigin AS, et al. FDG PET in the differential diagnosis of parkinsonian disorders. Neuroimage. 2005;26(3):912-21.

222. Osaki Y, Wenning GK, Daniel SE, Hughes A, Lees AJ, Mathias CJ, et al. Do published criteria improve clinical diagnostic accuracy in multiple system atrophy? Neurology. 2002;59(10):1486-91.

223. Osaki Y, Ben-Shlomo Y, Lees AJ, Daniel SE, Colosimo C, Wenning G, et al. Accuracy of clinical diagnosis of progressive supranuclear palsy. Mov Disord. 2004;19(2):181-9.

224. Vlaar AMM, van Kroonenburgh MJPG, Kessels AGH, Weber WEJ. Meta-analysis of the literature on diagnostic accuracy of SPECT in parkinsonian syndromes. BMC Neurology. 2007;7(27).

225. Kwon K-Y, Choi CG, Kim JS, Lee MC, Chung SJ. Comparison of brain MRI and 18F-FDG PET in the differential diagnosis of multiple system atrophy from Parkinson's disease. Mov Disord. 2007;22(16):2352-8.

226. National Institute for Health and Clinical Excellence - Social Care Institute for Excellence. Dementia. A NICE–SCIE Guideline on supporting people with dementia and their carers in health and social care. London: Social Care Institute for Excellence - National Institute for Health and Clinical Excellence; 2007. National Clinical Practice Guideline 42 Available from: http://www.scie.org.uk/publications/misc/dementia/dementia-fullguideline.pdf.

227. Swedish Council on Technology Assessment in Health Care. Dementia – Diagnostic and Therapeutic Interventions (vol 2). Systematic Review. SBU; 2008. (172E/2) Available from: http://www.sbu.se/upload/Publikationer/Content1/1/Dementia_Vol2.pdf.

228. Waldemar G, Dubois B, Emre M, Georges J, McKeith IG, Rossor M, et al. Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: EFNS guideline. Eur J Neurol. 2007;14(1):e1-26.

229. Yuan Y, Gu ZX, Wei WS. Fluorodeoxyglucose-positron-emission tomography, single-photon emission tomography, and structural MR imaging for prediction of rapid conversion to Alzheimer disease in patients with mild cognitive impairment: a meta-analysis. AJNR Am J Neuroradiol. 2009;30(2):404-10.

230. Whiting P, Gupta R, Burch J, Mujica Moto RE, Wright K, Marson A, et al. A systematic review of the effectiveness and cost-effectiveness of neuroimaging assessments used to visualise the seizure focus in people with refractory epilepsy being considered for surgery. Health Technology Assessment. 2006;10(4):iii-128.

231. Carne RP, O'Brien TJ, Kilpatrick CJ, MacGregor LR, Hicks RJ, Murphy MA, et al. MRI-negative PET-positive temporal lobe epilepsy: a distinct surgically remediable syndrome. Brain. 2004;127(Pt 10):2276-85.

232. Willmann O, Wennberg R, May T, Woermann FG, Pohlmann-Eden B. The contribution of 18F-FDG PET in preoperative epilepsy surgery evaluation for patients with temporal lobe epilepsy A meta-analysis. Seizure. 2007;16(6):509-20.

233. Medical Advisory Secretariat Ministry of Health and Long Term Care. Functional brain imaging: an evidence-based analysis. Ontario (Canada): MAS; 2006 December 2006. Ontario Heath Technology Assessment Series (6(22)) Available from: http://www.health.gov.on.ca/english/providers/program/mas/tech/reviews/pdf/rev_fbi_012507.pdf

234. Takaya S, Mikuni N, Mitsueda T, Satow T, Taki J, Kinoshita M, et al. Improved cerebral function in mesial temporal lobe epilepsy after subtemporal amygdalohippocampectomy. Brain. 2009;132(Pt 1):185-94.

235. Van Paesschen W, Dupont P, Sunaert S, Goffin K, Van Laere K. The use of SPECT and PET in routine clinical practice in epilepsy. Curr Opin Neurol. 2007;20(2):194-202.

236. Nelissen N, Van Paesschen W, Baete K, Van Laere K, Palmini A, Van Billoen H, et al. Correlations of interictal FDG-PET metabolism and ictal SPECT perfusion changes in human temporal lobe epilepsy with hippocampal sclerosis. Neuroimage. 2006;32(2):684-95.

237. Yun C-H, Lee SK, Lee SY, Kim KK, Jeong SW, Chung C-K. Prognostic factors in neocortical epilepsy surgery: multivariate analysis. Epilepsia. 2006;47(3):574-9.

238. Gaillard WD, Weinstein S, Conry J, Pearl PL, Fazilat S, Fazilat S, et al. Prognosis of children with partial epilepsy: MRI and serial 18FDG-PET. Neurology. 2007;68(9):655-9.

239. Tsakonas E, Moulton K, Spry C. FDG-PET to assess infections: a review of the evidence. Ottawa: Canadian Agency for Drugs and Technologies in Health (CADTH). 2008;Health Technology Assessment Rapid Review:23.

240. Termaat MF, Raijmakers PGHM, Scholten HJ, Bakker FC, Patka P, Haarman HJTM. The accuracy of diagnostic imaging for the assessment of chronic osteomyelitis: a systematic review and meta-analysis. J Bone Joint Surg Am. 2005;87(11):2464-71.

241. Kwee TC, Kwee RM, Alavi A. FDG-PET for diagnosing prosthetic joint infection: systematic review and metaanalysis. Eur J Nucl Med Mol Imaging. 2008;35(11):2122-32.

242. Keidar Z, Gurman-Balbir A, Gaitini D, Israel O. Fever of unknown origin: the role of 18F-FDG PET/CT. J Nucl Med. 2008;49(12):1980-5.

243. Keidar Z, Engel A, Hoffman A, Israel O, Nitecki S. Prosthetic vascular graft infection: the role of 18F-FDG PET/CT. J Nucl Med. 2007;48(8):1230-6.

244. Obyn C, Cleemput I, Léonard C, Closon J-P. Magnetic Resonance Imaging: cost analysis. Health Technology Assessment (HTA). Brussel: Belgian Health Care Knowledge Centre (KCE); 2009. KCE reports 106.

This page is left intentionally blank.

Dépôt légal : D/2009/10.273/25

KCE reports

1. Efficacité et rentabilité des thérapies de sevrage tabagique. D/2004/10.273/2.
2. Etude relative aux coûts potentiels liés à une éventuelle modification des règles du droit de la responsabilité médicale (Phase 1). D/2004/10.273/4.
3. Utilisation des antibiotiques en milieu hospitalier dans le cas de la pyélonéphrite aiguë. D/2004/10.273/6.
4. Leucoréduction. Une mesure envisageable dans le cadre de la politique nationale de sécurité des transfusions sanguines. D/2004/10.273/8.
5. Evaluation des risques préopératoires. D/2004/10.273/10.
6. Validation du rapport de la Commission d'examen du sous financement des hôpitaux. D/2004/10.273/12.
7. Recommandation nationale relative aux soins prénatals: Une base pour un itinéraire clinique de suivi de grossesses. D/2004/10.273/14.
8. Systèmes de financement des médicaments hospitaliers: étude descriptive de certains pays européens et du Canada. D/2004/10.273/16.
9. Feedback: évaluation de l'impact et des barrières à l'implémentation – Rapport de recherche: partie 1. D/2005/10.273/02.
10. Le coût des prothèses dentaires. D/2005/10.273/04.
11. Dépistage du cancer du sein. D/2005/10.273/06.
12. Etude d'une méthode de financement alternative pour le sang et les dérivés sanguins labiles dans les hôpitaux. D/2005/10.273/08.
13. Traitement endovasculaire de la sténose carotidienne. D/2005/10.273/10.
14. Variations des pratiques médicales hospitalières en cas d'infarctus aigu du myocarde en Belgique. D/2005/10.273/12
15. Evolution des dépenses de santé. D/2005/10.273/14.
16. Etude relative aux coûts potentiels liés à une éventuelle modification des règles du droit de la responsabilité médicale. Phase II : développement d'un modèle actuariel et premières estimations. D/2005/10.273/16.
17. Evaluation des montants de référence. D/2005/10.273/18.
18. Utilisation des itinéraires cliniques et guides de bonne pratique afin de déterminer de manière prospective les honoraires des médecins hospitaliers: plus facile à dire qu'à faire.. D/2005/10.273/20
19. Evaluation de l'impact d'une contribution personnelle forfaitaire sur le recours au service d'urgences. D/2005/10.273/22.
20. HTA Diagnostic Moléculaire en Belgique. D/2005/10.273/24, D/2005/10.273/26.
21. HTA Matériel de Stomie en Belgique. D/2005/10.273.28.
22. HTA Tomographie par Emission de Positrons en Belgique. D/2005/10.273/30.
23. HTA Le traitement électif endovasculaire de l'anévrysme de l'aorte abdominale (AAA). D/2005/10.273.33.
24. L'emploi des peptides natriurétiques dans l'approche diagnostique des patients présentant une suspicion de décompensation cardiaque. D/2005/10.273.35
25. Endoscopie par capsule. D2006/10.273.02.
26. Aspects médico-légaux des recommandations de bonne pratique médicale. D2006/10.273/06.
27. Qualité et organisation des soins du diabète de type 2. D2006/10.273/08.
28. Recommandations provisoires pour les évaluations pharmacoéconomiques en Belgique. D2006/10.273/11.
29. Recommandations nationales Collège d'oncologie : A. cadre général pour un manuel d'oncologie B. base scientifique pour itinéraires cliniques de diagnostic et traitement, cancer colorectal et cancer du testicule. D2006/10.273/13.
30. Inventaire des bases de données de soins de santé. D2006/10.273/15.
31. Health Technology Assessment : l'antigène prostatique spécifique (PSA) dans le dépistage du cancer de la prostate. D2006/10.273/18.
32. Feedback: évaluation de l'impact et des barrières à l'implémentation - Rapport de recherche: partie II. D2006/10.273/20.
33. Effets et coûts de la vaccination des enfants Belges au moyen du vaccin conjugué antipneumococcique. D2006/10.273/22.
34. Trastuzumab pour les stades précoces du cancer du sein. D2006/10.273/24.

35. Etude relative aux coûts potentiels liés à une éventuelle modification des règles du droit de la responsabilité médicale – Phase III : affinement des estimations. D2006/10.273/27.
36. Traitement pharmacologique et chirurgical de l'obésité. Prise en charge résidentielle des enfants sévèrement obèses en Belgique. D/2006/10.273/29.
37. Health Technology Assessment Imagerie par Résonance Magnétique. D/2006/10.273/33.
38. Dépistage du cancer du col de l'utérus et recherche du Papillomavirus humain (HPV). D/2006/10.273/36
39. Evaluation rapide de technologies émergentes s'appliquant à la colonne vertébrale : remplacement de disque intervertébral et vertébro/cyphoplastie par ballonnet. D/2006/10.273/39.
40. Etat fonctionnel du patient: un instrument potentiel pour le remboursement de la kinésithérapie en Belgique? D/2006/10.273/41.
41. Indicateurs de qualité cliniques. D/2006/10.273/44.
42. Etude des disparités de la chirurgie élective en Belgique. D/2006/10.273/46.
43. Mise à jour de recommandations de bonne pratique existantes. D/2006/10.273/49.
44. Procédure d'évaluation des dispositifs médicaux émergeants. D/2006/10.273/51.
45. HTA Dépistage du Cancer Colorectal : état des lieux scientifique et impact budgétaire pour la Belgique. D/2006/10.273/54.
46. Health Technology Assessment. Polysomnographie et monitoring à domicile des nourrissons en prévention de la mort subite. D/2006/10.273/60.
47. L'utilisation des médicaments dans les maisons de repos et les maisons de repos et de soins Belges. D/2006/10.273/62
48. Lombalgie chronique. D/2006/10.273/64.
49. Médicaments antiviraux en cas de grippe saisonnière et pandémique. Revue de littérature et recommandations de bonne pratique. D/2006/10.273/66.
50. Contributions personnelles en matière de soins de santé en Belgique. L'impact des suppléments. D/2006/10.273/69.
51. Besoin de soins chroniques des personnes âgées de 18 à 65 ans et atteintes de lésions cérébrales acquises. D/2007/10.273/02.
52. Rapid Assessment: Prévention cardiovasculaire primaire dans la pratique du médecin généraliste en Belgique. D/2007/10.273/04.
53. Financement des soins Infirmiers Hospitaliers. D/2007/10 273/06
54. Vaccination des nourrissons contre le rotavirus en Belgique. Analyse coût-efficacité
55. Valeur en termes de données probantes des informations écrites de l'industrie pharmaceutique destinées aux médecins généralistes. D/2007/10.273/13
56. Matériel orthopédique en Belgique: Health Technology Assessment. D/2007/10.273/15.
57. Organisation et Financement de la Réadaptation Locomotrice et Neurologique en Belgique D/2007/10.273/19
58. Le Défibrillateur Cardiaque Implantable.: un rapport d'évaluation de technologie de santé D/2007/10.273/22
59. Analyse de biologie clinique en médecine général. D/2007/10.273/25
60. Tests de la fonction pulmonaire chez l'adulte. D/2007/10.273/28
61. Traitement de plaies par pression négative: une évaluation rapide. D/2007/10.273/31
62. Radiothérapie Conformationelle avec Modulation d'intensité (IMRT). D/2007/10.273/33.
63. Support scientifique du Collège d'Oncologie: un guideline pour la prise en charge du cancer du sein. D/2007/10.273/36.
64. Vaccination HPV pour la prévention du cancer du col de l'utérus en Belgique: Health Technology Assessment. D/2007/10.273/42.
65. Organisation et financement du diagnostic génétique en Belgique. D/2007/10.273/45.
66. Drug Eluting Stents en Belgique: Health Technology Assessment. D/2007/10.273/48.
67. Hadronthérapie. D/2007/10.273/51.
68. Indemnisation des dommages résultant de soins de santé - Phase IV : Clé de répartition entre le Fonds et les assureurs. D/2007/10.273/53.
69. Assurance de Qualité pour le cancer du rectum – Phase I: Recommandation de bonne pratique pour la prise en charge du cancer rectal D/2007/10.273/55
70. Etude comparative des programmes d'accréditation hospitalière en Europe. D/2008/10.273/02
71. Recommandation de bonne pratique clinique pour cinq tests ophtalmiques. D/2008/10.273/05
72. L'offre de médecins en Belgique. Situation actuelle et défis. D/2008/10.273/08

73. Financement du programme de soins pour le patient gériatrique dans l'hôpital classique : Définition et évaluation du patient gériatrique, fonction de liaison et évaluation d'un instrument pour un financement approprié. D/2008/10.273/12
74. Oxygénothérapie Hyperbare: Rapid Assessment. D/2008/10.273/14.
75. Guideline pour la prise en charge du cancer oesophagien et gastrique: éléments scientifiques à destination du Collège d'Oncologie. D/2008/10.273/17.
76. Promotion de la qualité de la médecine générale en Belgique: status quo ou quo vadis ? D/2008/10.273/19.
77. Orthodontie chez les enfants et adolescents D/2008/10.273/21
78. Recommandations pour les évaluations pharmacoéconomiques en Belgique. D/2008/10.273/24.
79. Remboursement des radioisotopes en Belgique. D/2008/10.273/27.
80. Évaluation des effets du maximum à facturer sur la consommation et l'accessibilité financière des soins de santé. D/2008/10.273/36.
81. Assurance de qualité pour le cancer rectal – phase 2: développement et test d'un ensemble d'indicateurs de qualité. D/2008/10.273/39
82. Angiographie coronaire par tomodensitométrie 64-détecteurs chez les patients suspects de maladie coronarienne. D/2008/10.273/41
83. Comparaison internationale des règles de remboursement et aspects légaux de la chirurgie plastique D/2008/10.273/44
84. Les séjours psychiatriques de longue durée en lits T. D/2008/10.273/47
85. Comparaison de deux systèmes de financement des soins de première ligne en Belgique. D/2008/10.273/50.
86. Différenciation de fonctions dans les soins infirmiers :possibilités et limites D/2008/10.273/53
87. Consommation de kinésithérapie et de médecine physique et de réadaptation en Belgique. D/2008/10.273/55
88. Syndrome de Fatigue Chronique : diagnostic, traitement et organisation des soins. D/2008/10.273/59.
89. Evaluation des certains nouveaux traitements du cancer de la prostate et de l'hypertrophie bénigne de la prostate. D/2008/10.273/62
90. Médecine générale: comment promouvoir l'attraction et la rétention dans la profession ? D/2008/10.273/64.
91. Appareils auditifs en Belgique: health technology assessment. D/2008/10.273/68
92. Les infections nosocomiales en Belgique : Volet I, Etude Nationale de Prévalence. D/2008/10.273/71.
93. Détection des événements indésirables dans les bases de données administratives. D/2008/10.273/74.
94. Soins maternels intensifs (Maternal Intensive Care) en Belgique. D/2008/10.273/78.
95. Implantation percutanée des valvules cardiaques dans le cas de maladies valvulaires congénitales et dégénératives: A rapid Health Technology Assessment. D/2007/10.273/80.
96. Construction d'un index médical pour les contrats privés d'assurance maladie. D/2008/10.273/83.
97. Centres de réadaptation ORL/PSY : groupes cibles, preuves scientifiques et organisation des soins. D/2009/10.273/85.
98. Évaluation de programmes de vaccination généraux et ciblés contre l'hépatite A en Belgique. D/2008/10.273/89.
99. Financement de l'hôpital de jour gériatrique. D/2008/10.273/91.
100. Valeurs seuils pour le rapport coût-efficacité en soins de santé. D/2008/10.273/95.
101. Enregistrement vidéo des interventions chirurgicales par endoscopie : une évaluation rapide. D/2008/10.273/98.
102. Les infections nosocomiales en Belgique: Volet II: Impact sur la mortalité et sur les coûts. D/2009/10.273/100.
103. Réformes dans l'organisation des soins de santé mentale : étude d'évaluation des 'projets thérapeutiques' - 1er rapport intermédiaire. D/2009/10.273/05.
104. Chirurgie assistée par robot: health technology assessment. D/2009/10.273/08
105. Soutien scientifique au Collège d'Oncologie: recommandations pour la pratique clinique dans la prise en charge du cancer du pancréas. D/2009/10.273/11
106. Imagerie par résonance magnétique : analyse de coûts. D/2009/10.273/15
107. Indemnisation des dommages résultant de soins de santé. Phase V: impact budgétaire de la transposition du système français en Belgique. D/2009/10.273/17

108. Le Tiotropium dans le traitement des BronchoPneumopathies Chroniques Obstructives: Health Technology Assessment. D/2009/10.273/19
109. A propos de la valeur de l'EEG et des potentiels évoqués dans la pratique clinique. D/2009/10.273/22
110. La tomographie par émission de positrons en Belgique: une mise à jour. D/2009/10.273/25