DOI: 10.1111/1471-0528.12488
www.bjog.org

Systematic review

Evaluating PET-CT in the detection and management of recurrent cervical cancer: systematic reviews of diagnostic accuracy and subjective elicitation

C Meads,[a] C Davenport,[b] S Małysiak,[c] M Kowalska,[c] A Zapalska,[c] P Guest,[d] P Martin-Hirsch,[e] E Borowiack,[c] P Auguste,[f] P Barton,[f] T Roberts,[f] K Khan,[g] S Sundar[h]

[a] Health Economics Research Group, Brunel University, Middlesex, UK [b] Department of Public Health, Epidemiology and Biostatistics, University of Birmingham, Birmingham, UK [c] Arcana Institute, Krakow, Poland [d] Queen Elizabeth Medical Centre, University Hospitals Birmingham NHS Foundation Trust Queen Elizabeth Hospital, Birmingham, UK [e] Lancashire Teaching Hospitals NHS Trust, Royal Preston Hospital, Preston, UK [f] Health Economics Unit, School of Health and Population Sciences, University of Birmingham, Birmingham, UK [g] Centre for Primary Care and Public Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK [h] Pan Birmingham Gynaecological Cancer Centre, City Hospital and School of Cancer Sciences, University of Birmingham, Birmingham, UK

Correspondence: Dr S Sundar, Pan Birmingham Gynaecological Cancer Centre, City Hospital, Birmingham B18 7QH, UK.
Email s.s.sundar@bham.ac.uk

Accepted 9 September 2013. Published Online 3 December 2013.

Background Positron emission tomography–computed tomography (PET-CT) is recommended to triage women for exenterative surgery and surveillance after treatment for advanced cervical cancer.

Objective To evaluate diagnostic accuracy of additional whole body PET-CT compared with CT/magnetic resonance imaging (MRI) alone in women with suspected recurrent/persistent cervical cancer and in asymptomatic women as surveillance.

Design Systematic reviews. Subjective elicitation to supplement diagnostic information.

Search strategy/Selection criteria/Data collection and analysis Searches of electronic databases were performed to June 2013. Studies in women with suspected recurrent/persistent cervical cancer and in asymptomatic women undergoing follow up with sufficient numeric data were included. We calculated sensitivity, specificity and corresponding 95% confidence intervals. Meta-analyses employed a bivariate model that included a random-effects term for between-study variations (CT studies) and univariate random effects meta-analyses (PET-CT studies) for sensitivity and specificity separately.

Subjective elicitation Prevalence of recurrence and the accuracy of imaging elicited using the allocation of points technique. Coherence of elicited subjective probabilities with estimates in the literature examined.

Results We identified 15 relevant studies; none directly compared additional PET-CT with MRI or CT separately. Most CT and MRI studies used older protocols and the majority did not distinguish between asymptomatic and symptomatic women. Meta-analysis of nine PET-CT studies in mostly symptomatic women showed sensitivity of 94.8 (95% CI 91.2–96.9), and specificity of 86.9% (95% CI 82.2–90.5). The summary estimate of the sensitivity of CT for detection of recurrence was 89.64% (95% CI 81.59–94.41) and specificity was 76% (95% CI 43.68–92.82). Meta-analysis for MRI test accuracy studies was not possible because of clinical heterogeneity. The sensitivity and specificity of MRI in pelvic recurrence varied between 82 and 100% and between 78 and 100%, respectively. Formal statistical comparisons of the accuracy of index tests were not possible. Subjective elicitation provided estimates comparable to the literature. Subjective estimates of the increase in accuracy from the addition of PET-CT were less than elicited increases required to justify the use in PET-CT for surveillance.

Conclusion Evidence to support additional PET-CT is scarce, of average quality and does not distinguish between application for surveillance and diagnosis. Guidelines recommending PET-CT in recurrent cervical cancer need to be reconsidered in the light of the existing evidence base.

Keywords Accuracy, computed tomography, exenteration, magnetic resonance imaging, positron emission tomography-computed tomography, recurrent cervical cancer.

Please cite this paper as: Meads C, Davenport C, Małysiak S, Kowalska M, Zapalska A, Guest P, Martin-Hirsch P, Borowiack E, Auguste P, Barton P, Roberts T, Khan K, Sundar S. Evaluating PET-CT in the detection and management of recurrent cervical cancer: systematic reviews of diagnostic accuracy and subjective elicitation. BJOG 2013; DOI: 10.1111/1471-0528.12488.

Introduction

Cervical cancer was diagnosed in 2851 women in the UK in 2010 and 936 deaths from cervical cancer in the UK were recorded.[1] Early-stage cervical cancer is treated by surgery or chemoradiation (stages I–IIA) whereas advanced-stage cervical cancer (IIB–IIIB) is treated predominantly by chemoradiation. Chemotherapy alone is reserved for metastatic cancer at presentation. Recurrence is more common in advanced cervical cancer (30%) than in early-stage cervical cancer (6%).[2,3] Currently surveillance is based on clinical examination at regular follow-up visits to detect recurrence. If recurrence is suspected, either on the basis of symptoms or examination, computed tomography (CT) or magnetic resonance imaging (MRI) is used to confirm and define the extent of recurrence.[4] Neither modality can distinguish between radiation-induced fibrosis and malignancy.

Survival in women presenting with symptoms of recurrence—e.g. pain/bleeding/fistulae from locally advanced cancer or cachexia from distant metastases is substantially worse than in asymptomatic women detected at surveillance.[5–7] Treatment options for recurrent cervical cancer encompass radical surgery (salvage hysterectomy or pelvic exenteration), chemoradiotherapy and palliative treatment (which can be chemotherapy or radiotherapy).

In carefully selected women, with pelvis-confined or central recurrence, exenterative surgery involving the removal of bladder, uterus and vagina, and/or rectosigmoid is potentially curative. It is therefore reasonable to assume that improving early detection of recurrence in asymptomatic women will improve survival by identifying women with pelvis-confined recurrence where salvage surgery can be undertaken. However, salvage surgery carries the risk for significant morbidity and mortality, particularly where the pelvis has been irradiated. The long-term impact on the woman, including psychosocial effects, is also considerable. Accurately triaging women with distant metastases to receive palliative therapy and women with potentially curative central pelvic recurrence to exenterative surgery is critical to the management of women with recurrent cervical cancer.

Positron emission tomography (PET) uses ^{18}F-labelled fluorodeoxyglucose uptake in metabolically active tissues for the detection of malignancy. PET-CT combines PET with CT to define anatomical images. The CT images are used for localisation and characterisation of abnormal activity on the PET scans, and therefore they improve the specificity of the PET scan interpretation. However, registration CT scans performed as part of an integrated PET-CT study are almost universally carried out to a relatively low-dose protocol using lower exposure factors and thicker slices than dedicated diagnostic CT. Intravenous and oral contrast are not generally used. As a consequence, resolution and sensitivity for lesions for the registration CT alone will be lower than for a dedicated diagnostic CT.

Whole body PET-CT has shown promise in surveillance, (improved detection of recurrence and distant metastasis in asymptomatic women) and can predict survival outcome if performed 3 months after treatment.[8,9] However, PET-CT is expensive, the equipment alone costing about £2 million. False positives can occur in other metabolically active conditions, e.g. inflammation or sepsis.

PET-CT has been recently introduced into clinical practice to triage women for exenterative surgery and is endorsed for this use by national guidelines.[10,11] PET-CT is also recommended as surveillance after treatment for advanced-stage cervical cancer.[10] However, the diagnostic accuracy and effectiveness of PET-CT in accurately triaging women to potentially curative or palliative treatment and the diagnostic accuracy of PET-CT in asymptomatic women as surveillance for recurrence are not known. We performed systematic reviews of test accuracy and subjective elicitation to determine the diagnostic accuracy of whole body PET-CT in addition to CT/MRI in women following treatment for cervical cancer.

In identifying the additional value of PET-CT over standard CT/MRI, we sought evidence to answer three specific questions: (1) value of routine PET-CT in follow up of asymptomatic women after treatment for cervical cancer; (2) value of PET-CT imaging in detecting disease recurrence in symptomatic women; and (3) value of PET-CT imaging in recurrence to define a treatment strategy.

Methods

Systematic reviews of test accuracy

A generic protocol was developed for undertaking the systematic reviews of test accuracy, and diagnostic and therapeutic yield. Systematic reviews of test accuracy were conducted using established methods in line with the recommendations of the Cochrane Diagnostic Test Accuracy Working Group (http://srdta.cochrane.org/handbook-dta-reviews). Comprehensive searches from inception to June 2013 were conducted in MEDLINE, Embase, Science Citation Index, The Cochrane Library, MEDION, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, the Health Technology Assessment Database, Clinical Trials.com as well as a search of internet resources (UK Clinical Research Network Portfolio, specialist search gateways [OMNI and The National Cancer Institute], Google and Copernic). Electronic searches were supplemented by checking of reference lists, handsearching *Gynecologic Oncology* and contact with

authors of included studies. No language restrictions were applied.

Inclusion and exclusion criteria

Population

1 Included were:
- a any women with clinical suspicion of persistent or recurrent cervical cancer after primary treatment, on the basis of one or more of clinical history, clinical examination, tumour antigen tests and imaging.
- b any women who had had advanced stage cervical cancer (IB2–IV) treated previously, for example with chemoradiation, with a minimum gap between completion of treatment and imaging of 3 months and who were currently asymptomatic and undergoing routine follow up.

2 Excluded were:
- a studies where the population contained women within 3 months of completion of treatment for primary disease, due to problems associated with distinguishing treatment complications and inflammatory response from recurrence in this patient group.

Index test

1 Included was:
- a PET-CT using [^{18}F]fluorodeoxyglucose as radioisotope tracer in addition to routine imaging (CT/MRI).

2 Excluded was:
- a PET alone without concurrent CT.

Comparator tests

1 Included were:
- a CT, local or whole body.
- b MRI, local or whole body.
- c both CT and MRI, local or whole body.

Reference standard

1 Included were:
- a histopathological findings or clinical follow up for 6 months or more or both (differential verification: different reference standards for index test positive and index test negative women were accepted because of the difficulty of biopsy where there was no indicated lesion to biopsy in index test negative women).

2 Excluded were:
- a studies where only some of the participants undergoing the index test also received any reference standard (partial verification).

Outcomes

1 Included were:
- a studies that provided numerical data sufficient to create 2 × 2 tables of test results comparing index or comparator tests to the reference standard to provide information on test accuracy, giving true positive, true negative, false positive and false negative results.
- b studies that provided any information on diagnostic impact: change in diagnosis or staging after PET-CT compared with existing tests or reference standard.
- c studies that provided therapeutic impact: change in treatment plan after PET-CT compared with existing tests or reference standard.

Study design

1 Included were:
- a any prospective or retrospective test accuracy studies.
- b any diagnostic before and after studies investigating diagnostic and therapeutic impact with or without concurrent assessment of test accuracy.
- c studies with more than ten participants.

2 Excluded were:
- a studies on gynaecological cancers not providing separate data for the population with cervical cancer.
- b studies that described only lesion-based analysis rather than person-based analysis.

Study selection, data extraction and quality assessment

Inclusion of studies, data extraction and quality assessment were carried out in duplicate using predesigned and piloted data extraction forms and the QUADAS quality assessment tool for evaluations of test accuracy.[12] Differences were resolved by consensus and/or arbitration involving a third reviewer. Information on the technical quality of imaging technologies were also collected.

Statistical analysis

Data were extracted as two-by-two tables (true positives, false positives, true negatives and false negatives) and REVMAN version 5.2 and STATA version 11 were used for analysis. Where equivocal results were reported these were used in sensitivity analyses by adding the total number of equivocal results to each of true positives, false positives, false negatives and true negatives in turn to derive maximum and minimum variation in sensitivity and specificity. Results were displayed graphically on Forest and receiver operator characteristics (ROC) plots.[13] A bivariate model that included a random-effects term for variation in accuracy and threshold between studies was fitted where adequate results were available to derive summary estimates of sensitivity and specificity and summary ROC curves. The bivariate model has two levels corresponding to variation within and between studies in the meta-analysis.[14] At the first level, the within-study variability for both sensitivity and specificity is assumed to

follow a binomial distribution. The sensitivity–specificity pair for each study must be modelled jointly within study at level one of the analysis because they are correlated. At the second level variation between studies is modelled.[15] Where the model failed to converge or a correlation could not be estimated properly the bivariate model was simplified to two univariate random effects logistic regression models.

Methods for subjective elicitation

An elicitation exercise with specialists in gynaecological imaging, radiation oncology and gynaecological oncology was planned in anticipation of a lack of evidence with which to undertake an economic analysis[16]; in particular disaggregation of estimates of prevalence and test performance in asymptomatic and symptomatic women and direct comparisons of testing strategies (CT and/or MRI versus routine addition of PET-CT to CT and/or MRI).[17]

The subjective elicitation exercise aimed to answer (1) the accuracy of routine imaging (CT/MRI); (2) the accuracy of routine PET-CT in follow-up of asymptomatic women after treatment for cervical cancer; (3) the accuracy of the addition of PET-CT imaging to CT/MRI in detecting a recurrence in symptomatic women; and (4) the incremental accuracy required to justify the addition of PET-CT to CT/MRI in routine practice.

Probabilities elicited

Informed by the preliminary results of the systematic reviews of test accuracy, the research team decided on the data priorities for elicitation as follows:

1 To determine the prevalence of recurrence in women with an initial diagnosis of stage IB–IVA cervical cancer, where women are assumed to be disease free for a minimum of 3 months after completion of primary treatment: in women presenting with symptoms suggestive of recurrence and separately in asymptomatic women.
2 To determine the test accuracy of chest, abdominal and pelvic CT and/or MRI performed at the discretion of clinicians in women with an initial diagnosis of stage IB–IVA cervical cancer, who are assumed to be disease free for a minimum of 3 months after completion of primary treatment: in women presenting with symptoms suggestive of recurrence, and in asymptomatic women (CT/MRI as surveillance).
3 To determine the test accuracy of CT and/or MRI performed at the discretion of clinicians and of PET-CT (performed regardless of the result of initial imaging) in women with an initial diagnosis of stage IB–IVA cervical cancer, who are assumed to be disease free a minimum of 3 months after completion of primary treatment: in women presenting with symptoms suggestive of recurrence, and in asymptomatic women (CT and/or MRI + PET-CT used for surveillance).

The initial elicitation exercise ($n = 9$) was facilitated during an educational meeting to evaluate the accessibility of materials for respondents. Following the success of the initial elicitation, as judged by the face validity of findings fed back to participants, elicitations from subsequent specialists ($n = 12$) were conducted using self-completed questionnaires. Subjective estimates of the prevalence of cervical cancer recurrence in two hypothetical cohorts of symptomatic and asymptomatic women and the accuracy of two testing strategies (CT and/or MRI performed at the discretion of clinicians and the routine addition of PET-CT performed regardless of the result of CT and/or MRI) were elicited. Participants completed the elicitation exercise independently in order to ensure that any variation within and across disciplines could be captured. The elicitation exercise comprised an 11-page anonymous self-administered questionnaire (see Appendix S1). We collected data on experience, use of current imaging techniques and participant's use of PET-CT. We asked what participants considered to be the minimum important clinical difference (in terms of test error rates) in accuracy between imaging with CT and/or MRI alone compared with routine addition of PET-CT to CT and/or MRI that they would require before the introduction of one or other imaging strategy into practice.

Accuracy data were elicited in the form of the proportion of test errors (false positives and false negatives). We chose test errors as a metric of accuracy based on research suggesting that the clinical utility of a test is commonly conceptualised in this way.[14,18] Subjective estimates of test error rates and of the prevalence of cervical cancer recurrence were used to derive positive predictive values (PPV) and negative predictive values (NPV) for asymptomatic and symptomatic women separately.

We defined PPV as the proportion of women who test positive on CT and/or MRI (and separately if PET-CT was to be added routinely) who are confirmed as having recurrence of disease on the basis of histology. NPV is defined as the proportion of women who test negative on CT and/or MRI (and separately for routine addition of PET-CT) who are confirmed as not having recurrence on the basis of a minimum of 6 months clinical follow up. Elicitation of prevalence and test accuracy information was undertaken using the allocation of points technique whereby respondents are asked to indicate the likelihood of a value range being a true estimate by allocating a proportion of 100 points to that value range (the sum of allocated points across each value range summing to 100). Value ranges differed depending on the question being asked. For example the spread of value ranges for subjective estimates of the prevalence of recurrence in asymptomatic women was 0–49% including a single category for >50%. The spread of

value ranges for subjective estimates of the prevalence of recurrence in symptomatic women was 51–100% including a single category of <50%. For elicitation of test accuracy (false positives and false negatives) the spread of value ranges was between 0 and 50% to reflect the fact that a test error rate greater than 50% equates to a test accuracy that is worse than chance. In this way probability functions were obtained for each individual and aggregated mathematically to derive an average distribution for the sample.[18] An aggregated mean value was estimated using the average distribution and the mid-point of each value range. The variability of this aggregated mean was estimated by calculating the standard deviation across the value ranges.

Results

Results of systematic review of test accuracy

Study selection and characteristics of included studies

From 7719 potentially relevant citations, we selected 261 full-text articles for assessment. A total of 246 articles were excluded, most often for different patient population or incorrect study design. Figure 1 shows the PRISMA diagram of selection process. Nine studies evaluated PET-CT,[19–27] two evaluated MRI,[28,29] four evaluated CT,[20,30–32] one evaluated both MRI and CT.[33] Three studies gave results for both CT and/or MRI versus CT and/or MRI with whole body PET-CT with the same reference standard of histology or clinical evidence of disease in one table so comparisons can be drawn.[20,22,25] Tables S1, S2 and S3 describe the characteristics of included women. The total number of women in the studies ranged from 20 to 276 but some of the studies included women with any gynaecological cancers and others reported imaging results for both recurrent and primary cervical cancer.

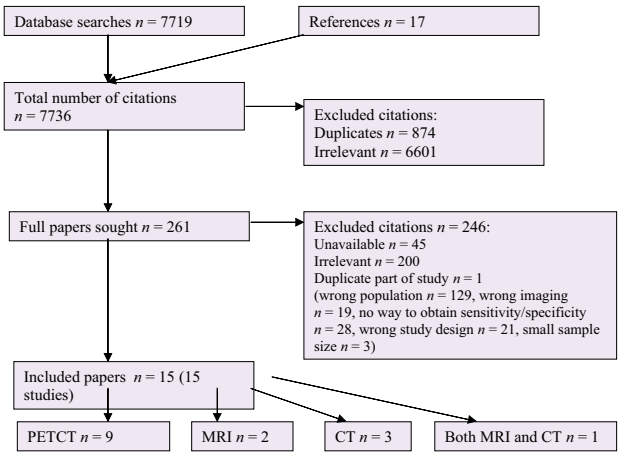

Figure 1. PRISMA diagram of systematic review.

Of note, the majority of CT/MRI studies were published between 1981 and 2000 and most did not use standard imaging methods. The quality of the studies was poor; in particular very little clinical information about participants was given and incorporation bias was inevitable for index-test-negative women as a result of the reference standard being clinical follow up, which is likely to have included imaging (see Table S4).

Included studies for each study question

1 Value of routine PET-CT in follow-up of women after treatment for cervical cancer. Two studies included asymptomatic women. One did not present data for asymptomatic women separately from women with symptoms.[24] One study gave recurrence rates for routine surveillance and suspected recurrence separately, but not mortality rates in each group.[21]
2 Value of PET-CT imaging in detecting a recurrence in case of symptoms. We found nine relevant studies.[19–27]
3 Value of PET-CT imaging in order to define the treatment strategy. Two included PET-CT studies reported information on diagnostic and therapeutic impact.[21,25]

Statistical results for accuracy of imaging

The sensitivity and specificity for detection of cervical cancer recurrence with CT ranged between 78 and 93% and 50 and 100%, respectively. The summary estimate of the sensitivity of CT for detection of recurrence based on the bivariate hierarchical model was 89.64 (95% CI 81.59–94.41) and specificity was 76 (95% CI 43.68–92.82).

The bivariate model failed to converge for estimation of the accuracy of PET-CT and estiamtes are therefore based on univariate random effects meta-analyses for sensitivity and specificity separately. This approach was considered appropriate because the failure of the model to converge appeared to be due to a lack of correlation between sensitivity and specificity (see Figure 2, forest plot sensitivity and specificity PET-CT). The sensitivities and specificities of detection of local and distant recurrence combined with PET-CT ranged between 83 and 100% and between 50 and 100%, respectively. For distant recurrence alone the sensitivity of PET-CT was 86% and the specificity was 100%. The summary estimate of the sensitivity of PET-CT for detection of cervical cancer recurrence was 94.8% (95% CI 91.2–96.9) and specificity was 86.9 (95% CI 82.2–90.5) (Figure 2). A sensitivity analysis omitting one study (Amit et al.[19]) that reported accuracy for distant recurrence only did not affect accuracy estimates to any significant degree (sensitivity 95.0, 95% CI 91.4–97.2; specificity 86.7, 95% CI 81.9–90.4). Only three studies ($n = 15$) gave results for both standard imaging alone and standard imaging with the addition of whole body PET-CT with the same reference standard of histology or clinical evidence of disease in

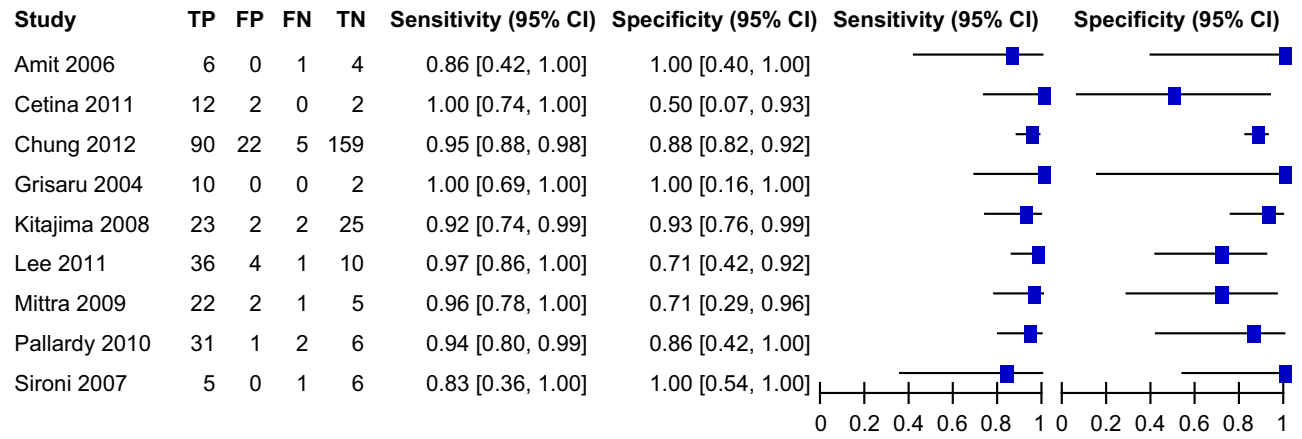

Figure 2. Sensitivity and specificity of PET-CT studies.

one table. Unfortunately the part of the body imaged with standard imaging was not mentioned in one paper (Grisaru).[22] This demonstrated sensitivity and specificity of CT and/or MRI of 25 and 50%, respectively, while the addition of PET-CT to this imaging strategy resulted in a sensitivity and specificity of 100%.[22] In Cetina et al.[20] the sensitivity of CT alone was 91.7%, which improved to 100% with the addition of PET-CT. There was no difference in specificity between the two imaging strategies (50%). Figure 3 shows Forest plot of sensitivity and specificity of CT studies. In Pallardy et al.[25] the sensitivity of routine imaging alone compared with routine imaging with the addition of PET-CT was 43% compared with 94% and the corresponding values for specificity were 72 and 75%. Numbers of women imaged in all three studies were small.

Meta-analysis for MRI test accuracy studies was not possible because of considerable clinical heterogeneity. The sensitivity and specificity of MRI in pelvic recurrence varied between 82 and 100% and between 78 and 100%, respectively.

Results for subjective elicitation

Subjective estimation of the prevalence of recurrence was elicited from all 21 respondents and subjective estimation of accuracy from 18 respondents. Responses from individuals who received pre-elicitation education in the form of a lecture did not appear to differ from those completing self-administered questionnaires only. The self-reported characteristics of respondents and their reported use of imaging technologies are outlined in Table S5 and in Figure 4. The mean elicited prevalence of recurrence in women presenting with symptoms a minimum of 3 months after completion of primary treatment was 47.8% (SD 20.8) and for asymptomatic women was 16.7% (SD 13.1). Subjective estimates of the accuracy of the two testing strategies and the minimum important difference between them considered sufficient to warrant the routine addition of PET-CT for the detection of cervical cancer recurrence are shown in Table 1. Mean elicited estimates of the increase in PPV of CT and/or MRI plus PET-CT compared with CT and/or MRI alone in symptomatic women was 2.6% and the increase in NPV was 3.6%. For asymptomatic women the mean elicited increase in PPV was 4.6% and in NPV 3.4%.

The minimum important elicited increase in accuracy of the addition of PET-CT to CT and/or MRI considered necessary to warrant the introduction of PET-CT as a routine investigation in this sample of clinical experts was similar for asymptomatic women (a mean 8.7% reduction in false positives and 6.3% reduction in false negatives) and

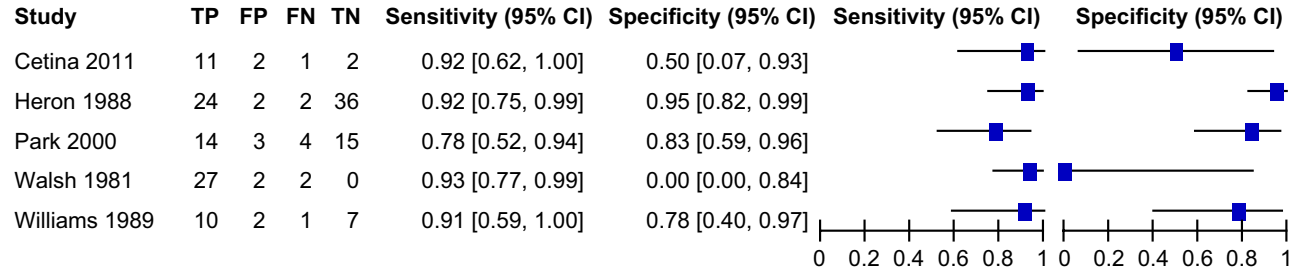

Figure 3. Sensitivity and specificity of CT studies.

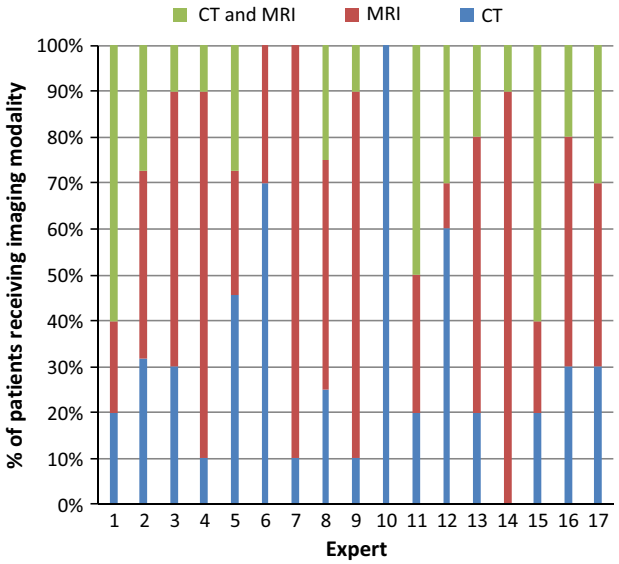

Figure 4. Use of imaging (MRI and/or CT in women presenting with suspected cervical cancer recurrence.

Table 1. Summary of accuracy results from subjective elicitation exercise

	MRI and/or CT mean (SD)	MRI and/or CT and PET-CT mean (SD)	Difference in false positives and in false negatives
Symptomatic			
PPV	88.4 (9.2)	91.0 (8.2)	2.6
NPV	86.8 (8.7)	90.7 (7.2)	3.6
Asymptomatic			
PPV	85.6 (9.8)	90.2 (7.7)	4.6
NPV	90.0 (7.7)	93.4 (5.5)	3.4

symptomatic women (a mean 7.7% reduction in false positives and 6.4% reduction in false negatives). Hence the subjective estimate of incremental accuracy resulting from the routine addition of PET-CT to MRI and/or CT was estimated to be smaller than the elicited minimum important difference in accuracy required to justify its use for the investigation of women after completion of primary treatment for cervical cancer.

Comparison with systematic review results

We found that elicited estimates of the accuracy of CT and/or MRI plus PET-CT compared with CT and/or MRI alone in symptomatic women were similar to estimates of accuracy in the literature (Table 2). The absence of published estimates of accuracy in asymptomatic women precluded a comparison in this group. Elicited specificities of CT and/or MRI and CT and/or MRI plus PET-CT in asymptomatic women were comparable to literature-based estimates in mixed symptomatic and asymptomatic populations whereas elicited sensitivities were lower. A lower sensitivity would be expected in a homogeneous asymptomatic population compared with a mixed symptomatic and asymptomatic population and therefore this finding supports the validity of elicited estimates.

Discussion

Main findings

Intercollegiate guidelines recommend the use of PET-CT in women with recurrent cervical cancer considered for exenterative surgery or where previous imaging is equivocal. The evidence underpinning these recommendations was largely derived from studies of diagnostic accuracy of PET-CT in primary cervical cancer to predict lymph node metastasis.[34–36] In addition, SIGN guidelines also

Table 2. Comparison of test accuracy results from elicitation exercise and systematic review of literature

Presentation	Asymptomatic				Symptomatic			
	Literature		Elicited***		Literature		Elicited***	
Accuracy	Sens	Spec	Sens	Spec	Sens	Spec	Sens	Spec
Clinical follow up and MRI ± CT	–	–	45.43	98.47			85.09	89.78
CT	–	–	–	–	78–93*	78–95**	–	–
MRI	–	–	–	–	82–100*	78–100**	–	–
Clinical follow up, MRI ± CT and PET-CT	–	–	65.25	98.58	83–100*	71–100**	89.71	91.88

Estimates of *sensitivity and **specificity for CT, MRI and PET-CT based mainly on symptomatic but frequently not distinguished according to presentation (asymptomatic or symptomatic) in the literature.
***Elicited estimates of sensitivity and specificity based on prevalence of recurrence in asymptomatic women of 16.7% and in symptomatic women of 47.8%.

recommend a PET-CT scan 9 months after chemoradiotherapy based on limited evidence.[37,38]

We evaluated evidence to answer three relevant questions: value of PET-CT in routine follow up in asymptomatic women, the value of PET-CT in women with symptoms suspicious of recurrence and the value of PET-CT in defining therapy. In particular, we sought to identify the additional value of PET-CT over conventional CT/MRI in these clinical scenarios. Our systematic review finds that evidence of diagnostic accuracy to support the use of whole body PET-CT in addition to standard CT or MRI in all three scenarios is scarce and of average quality. We found that published studies often do not distinguish between applications for surveillance and for diagnosis in suspected recurrence. There was scant information on imaging as routine follow up for asymptomatic women. Only two papers on diagnostic impact were found.[21,25] In particular most of the MRI and CT studies did not reflect current practice standards so the true additional value of PET-CT in these scenarios was unclear. In fact, most included PET-CT studies present results of diagnostic accuracy of PET-CT alone rather than the accuracy of PET-CT in comparison with CT/MRI. Hence the additional value of PET-CT in these settings is unclear. Only meta-analysis of PET-CT results was possible and results from the literature were coherent with findings of the subjective elicitation exercise.

The elicited estimated increase in accuracy of adding PET-CT to MRI and/or CT was less than the elicited minimum important difference in accuracy required to justify the routine addition of PET-CT for the investigation of women after completion of primary treatment for cervical cancer.

Strengths and limitations

Our systematic review was comprehensive in its scope and search. We conducted the review in line with contemporary recommendations. Our search of the literature aimed to minimise the risk of selection and publication bias. We made considerable efforts to find appropriate input values on effectiveness of treatment for the decision analytic model, which is based on best available evidence. All assumptions used in the model were agreed by the team based on expert advice a priori.

Experts used in the elicitation exercise were representative in specialty and expertise of decision makers in recurrent cervical cancer. The subjective elicitation exercise was carried out using expert opinion, before any economic analysis was undertaken and produced data not available in the published literature. The definition of expert for the purposes of subjective elicitation is not considered restricted to hands-on experience of a technology as subjective beliefs are shaped by factors other than first-hand experience such as interaction with colleagues, published estimates of accuracy and knowledge of the technology.[18]

Elicited estimates of accuracy of CT, MRI and PET-CT are plausible and reflect the fact that the accuracy of imaging tests is likely to be greater in symptomatic women than asymptomatic women. In addition, the pattern of elicited estimates of accuracy in asymptomatic women is plausible given the lower prevalence of recurrence in this group. Elicited estimates of accuracy also reflect a greater likelihood of an improvement in NPV compared with PPV in both symptomatic and asymptomatic women, which is consistent with the probability of a larger number of false positives with the addition of PET-CT to current imaging practice. Importantly, elicited estimates of prevalence and accuracy had face validity as judged by feedback to clinical experts who participated in the face to face elicitation exercise.

We did not evaluate subjective estimates of the accuracy of the use of CT/MRI to triage for subsequent testing with PET-CT, which is recommended practice.[10,11] However, the systematic searches did not identify any papers to support a selective approach either. This work is part of a larger National Institute of Health Research/Health Technology Assessment (HTA reference 09/29/02) funded evaluation of the accuracy and cost effectiveness of PET-CT in recurrent cervical cancer.[17] We identified the best inputs from evidence to construct a decision analytic model to determine the cost effectiveness of additional PET-CT in recurrent cervical cancer.

Interpretation

Current guidelines support the use of PET-CT in suspected recurrent or persistent disease after initial imaging and as surveillance in asymptomatic women after completion of chemoradiation for primary treatment. This is not supported by evidence from this systematic review of the literature or by the elicitation exercise. Published literature does not support The use of PET-CT for surveillance in asymptomatic patients after completion of primary treatment and is unclear on the value of additional PET-CT in selecting patients for exenterative surgery. This research finds that the accuracy of PET-CT in recurrent cervical cancer is not yet proven. However the authors acknowledge that lack of evidence of the value (of PET-CT) is not the same as evidence to support lack of value.

Good quality, adequately powered studies directly comparing the test accuracy of the addition of PET-CT to MRI and/or CT imaging alone in women with recurrent and persistent cervical cancer are needed. Studies also need to investigate the impact of additional PET-CT on change in diagnosis, work-up and change in the treatment plan. We also recommend that a national register of women considered for exenterative surgery for recurrent cervical cancer be established to prospectively collect data on imaging, decision making and outcomes of treatment. To our

knowledge, although elicitation of conditional probabilities about treatment effects have been undertaken previously, this is the first specific example of elicitation of test accuracy estimates of which we are aware and it demonstrates the value of this approach to inform subsequent modelling where primary data are scant or unreliable.[18] Further test accuracy elicitation exercises will be required to confirm the validity of this approach and for comparison of test accuracy elicitation using other test accuracy metrics. Investigation of the benefit of face to face pre-elicitation education on the validity of responses is warranted as this has an impact on the methods of elicitation that are possible (for example the use of postal and internet-based questionnaires), the resources required and response rate.

Conclusions

Our study raises two important issues—first the paucity of robust evidence on which to base decision making in the diagnosis and treatment of recurrent cervical cancer, and second a broader question on how rapidly evolving, often 'glamorous' technology can be robustly evaluated before its incorporation into routine clinical care. The use of PET-CT in recurrent cervical cancer and its endorsement by national guidelines is not currently supported by published literature. Consideration to revise national guidelines and/or prospective study in a national registry of exenterative surgery for recurrent cervical cancer to evaluate the effectiveness of PET-CT in this setting is necessary.

Disclosure of interest

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf and declare that (1) no authors have support from any company for the submitted work; (2) no authors have any relationships from any company that might have an interest in the submitted work in the previous 3 years; (3) their spouses, partners, or children have no financial relationships that may be relevant to the submitted work; and (4) no authors have any nonfinancial interests that may be relevant to the submitted work.

Contribution to authorship

SS, KK, TER contributed to the design of the whole project and obtained funding. CM and CD designed and supervised the systematic review with input from SS, TER, PB, PA, PM-H and PG. The systematic searches and meta-analysis were performed by AZ, SM, MK, AZ and EB. CM and CD designed and conducted the subjective elicitation with input from SS, PM-H, TER, PB, PG and PA. SS and CM prepared the manuscript as lead writers. All authors contributed to critical review and input into the final manuscript. CM is the guarantor.

Details of ethics approval

No ethics approval was required for this systematic review and subjective elicitation. Subjective elicitation was performed with written consent from participants.

Funding

This review was funded by the National Institute for Health Research Health Technology Assessment Programme (09/29/02).

Acknowledgements

The authors gratefully acknowledge the following: Shakila Thangaratinam for drafting the grant application, Anne Fry-Smith for helping to draft the original protocol, Paweł Chomiak for assistance with the systematic reviews and Celia Taylor for help with the analysis strategy for the elicitation exercise. The following are thanked for participating in the subjective elicitation: Dr Indrajit Fernando consultant clinical oncologist, Mr KK Chan consultant gynaecological oncology, Mr Janos Balega consultant gynaecological oncology, Ahmed Elattar subspecialty trainee in gynaecological oncology, Raj Saha consultant gynaecologist (unit lead), Martin Underwood specialist trainee obstetrics and gynaecology, Dr Moji Balogun consultant radiologist, Prof David Luesley consultant gynaecologist (Director Pan Birmingham Gynaecological Cancer Centre), Raj Naik consultant gynaecological oncology, Adam Rosenthal consultant gynaecological oncology, Arjun Jeyarajah consultant gynaecological oncology, John Butler subspecialty trainee in gynaecological oncology, Alex Lawrence consultant gynaecological oncology, NJ Wood consultant gynaecological oncology, Mary Cairns consultant gynaecological oncology, John Shepherd consultant surgeon/gynaecological oncology, Miss Elizabeth Ball, consultant obstetrician and gynaecologist.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Studies included in the diagnostic review.

Table S2. Definitions of reference standards presented in included studies.

Table S3. Population characteristics of studies evaluating PET-CT.

Table S4. Quality of included studies.

Table S5. Characteristics of respondents to the elicitation exercise.

Appendix S1. Questionnaire for subjective elicitation.

References

1 Cancer Research UK. 2012 [www.cancerresearchuk.org/cancer-info/cancerstats/types/cervix/]. Accessed 27 October 2013.

2 Kim MK, Jo H, Kong HJ, Kim HC, Kim JW, Kim YM, et al. Postoperative nomogram predicting risk of recurrence after radical hysterectomy for early-stage cervical cancer. *Int J Gynecol Cancer* 2010;20:1581–6.

3 Hirakawa M, Nagai Y, Inamine M, Ogawa K, Toita T, Murayama S, et al. Predictive factors of distant recurrence in locally advanced squamous cell carcinoma of the cervix treated with concurrent chemoradiotherapy. *Gynecol Oncol* 2008;108:126–9.

4 Anon. Topotecan (Hycamtin©) for the treatment of recurrent and stage IVB carcinoma of the cervix. *Single Technology Appraisal (STA) Submission to the National Institute for Health and Clinical Excellence*. London: GlaxoSmithKline, 2009.

5 Bodurka-Bevers D, Morris M, Eifel PJ, Levenback C, Bevers MW, Lucas KR, et al. Posttherapy surveillance of women with cervical cancer: an outcomes analysis. *Gynecol Oncol* 2000;78:187–93.

6 Lim KC, Howells RE, Evans AS. The role of clinical follow up in early stage cervical cancer in South Wales. *BJOG* 2004;111:1444–8.

7 Brooks RA, Rader JS, Dehdashti F, Mutch DG, Powell MA, Thaker PH, et al. Surveillance FDG-PET detection of asymptomatic recurrences in patients with cervical cancer. *Gynecol Oncol* 2009;112:104–9.

8 Chung HH, Jo H, Kang WJ, Kim JW, Park NH, Song YS, et al. Clinical impact of integrated PET/CT on the management of suspected cervical cancer recurrence. *Gynecol Oncol* 2007;104:529–34.

9 Jover R, Lourido D, Gonzalez C, Rojo A, Gorospe L, Alfonso JM. Role of PET/CT in the evaluation of cervical cancer. *Gynecol Oncol* 2008;110 (3 Suppl 2):S55–9.

10 Scottish Intercollegiate Guidelines N. *Management of Cervical Cancer*. Edinburgh: NHS Quality Improvement Scotland, 2008.

11 Barrington S, Scarsbrook A. *Evidence Based Indications for the Use of PET-CT in the United Kingdom 2012*. London: The Royal College of Physicians and The Royal College of Radiologists, 2012.

12 Diagnostic Test Accuracy Working G. *Handbook for DTA Reviews*. 2012 [http://srdta.cochrane.org/handbook-dta-reviews]. Accessed 27 October 2013.

13 Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2010;3:25.

14 Reitsma JB, Glas AS, Rutges AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982–90.

15 Meads CA, Davenport CF. Quality assessment of diagnostic before-after studies: development of methodology in the context of a systematic review. *BMC Med Res Methodol* 2009;9:3.

16 Auguste P, Barton P, Meads C, Davenport C, Małysiak S, Kowalska M, et al. Evaluating PET–CT in routine surveillance and follow-up after treatment for cervical cancer: a cost-effectiveness analysis. *BJOG*. doi: 10.1111/1471-0528.12460 (forthcoming).

17 Meads C, Auguste P, Davenport C, Malysiak S, Sundar S, Kowalska M, et al. Positron emission tomography/computerised tomography imaging in detecting and managing recurrent cervical cancer: systematic review of evidence, elicitation of subjective probabilities and economic modelling. *Health Technol Assess* 2013;17:1–323.

18 O'Hagan A, Buck CE, Daneschkkah A, Eiser JR, Garthwaite PH, Jenkinson DJ, et al. Uncertain judgements: eliciting expert's probabilties. *Stat Pract* 2006.

19 Amit A, Beck D, Lowenstein L, Lavie O, Bar S, Kedar Z, et al. The role of hybrid PET/CT in the evaluation of patients with cervical cancer. *Gynecol Oncol* 2006;100:65–9.

20 Cetina L, Serrano A, Cantu-de-Leon D, Perez-Montiel D, Estrada E, Coronel J, et al. F18-FDG-PET/CT in the evaluation of patients with suspected recurrent or persistent locally advanced cervical carcinoma. *Rev Invest Clin* 2011;63:227–35.

21 Chung HH, Kim JW, Kang KW, Park NH, Song YS, Chung JK, et al. Predictive role of post-treatment [^{18}F]FDG PET/CT in patients with uterine cervical cancer. *Eur J Radiol* 2012;81:e817–22.

22 Grisaru D, Almog B, Levine C, Metser U, Fishman A, Lerman H, et al. The diagnostic accuracy of ^{18}F-fluorodeoxyglucose PET/CT in patients with gynecological malignancies. *Gynecol Oncol* 2004;94:680–4.

23 Kitajima K, Murakami K, Yamasaki E, Domeki Y, Kaji Y, Sugimura K. Performance of FDG-PET-CT for diagnosis of recurrent uterine cervical cancer. *Eur Radiol* 2008;18:2040–7.

24 Mittra E, El-Maghraby T, Rodriguez CA, Quon A, Ross McDougal I, Gambhir SS, et al. Efficacy of ^{18}F-FDG PET-CT in the evaluation of patients with recurrent cervical carcinoma. *Eur J Nucl Med Mol Imaging* 2009;36:1952–9.

25 Pallardy A, Bodet-Milin C, Oudoux A, Campion L, Bourbouloux E, Sagan C, et al. Clinical and survival impact of FDG PET in patients with suspicion of recurrent cervical carcinoma. *Eur J Nucl Med Mol Imaging* 2010;37:1270–8.

26 Sironi S, Picchio M, Landoni C, Galimberti S, Signorelli M, Bettinardi V, et al. Post-therapy surveillance of patients with uterine cancers: value of integrated FDG PET/CT in the detection of recurrence. *Eur J Nucl Med Mol Imaging* 2007;34:472–9.

27 Minkyung Lee YL, Hwang KH, Choe W, Park CY. Usefulness of F-18 FDG PET/CT in assessment of recurrence of cervical cancer after treatment. *Nucl Med Mol Imaging* 2011;45:111–6.

28 Weber TM, Sostman HD, Spritzer CE, Ballard RL, Meyer GA, Clark-Pearson DL, et al. Cervical carcinoma: determination of recurrent tumor extent versus radiation changes with MR imaging. *Radiology* 1995;194:135–9.

29 Hatano K. Evaluation of the therapeutic effect of radiotherapy on cervical cancer using magnetic resonance imaging. *Int J Radiat Oncol Biol Phys* 1999;45:639–44.

30 Park DH, Kim KH, Park SY, Lee BH, Choi CW, Chin SY. Diagnosis of recurrent uterine cervical cancer: computed tomography versus positron emission tomography. *Korean J Radiol* 2000;1:51–5.

31 Heron CW, Husband JE, Williams MP, Dobbs J, Cosgrove DO. The value of CT in the diagnosis of recurrent carcinoma of the cervix. *Clin Radiol* 1988;39:501.

32 Walsh JW, Amendola MA, Hall DJ, Tisnado J, Goplerud DR. Recurrent carcinoma of the cervix: CT diagnosis. *Am J Roentgenol* 1981;136:117–22.

33 Williams MP, Husband JE, Heron CW, Cherryman GR, Koslin DB. Magnetic resonance imaging in recurrent carcinoma of the cervix. *Br J Radiol* 1989;62:544–50.

34 Loft A, Berthelsen AK, Roed H, Ottosen C, Lundvall L, Knudsen J, et al. The diagnostic value of PET/CT scanning in patients with cervical cancer: a prospective study. *Gynecol Oncol* 2007;106:29–34.

35 Boughanim M, Leboulleux S, Rey A, Pham CT, Zafrani Y, Duvillard P, et al. Histologic results of para-aortic lymphadenectomy in patients treated for stage IB2/II cervical cancer with negative [^{18}F] fluorodeoxyglucose positron emission tomography scans in the para-aortic area. *J Clin Oncol* 2008;26:2558–61.

36 Chao A, Ho KC, Wang CC, Cheng HH, Lin G, Yen TC, et al. Positron emission tomography in evaluating the feasibility of curative intent in cervical cancer patients with limited distant lymph node metastases. *Gynecol Oncol* 2008;110:172–8.

37 Grigsby PW, Siegel BA, Dehdashti F, Rader J, Zoberi I. Posttherapy [^{18}F]fluorodeoxyglucose positron emission tomography in carcinoma of the cervix: response and outcome. *J Clin Oncol* 2004;22:2167–71.

38 Ryu SY, Kim MH, Choi SC, Choi CW, Lee KH. Detection of early recurrence with ^{18}F-FDG PET in patients with cervical cancer. *J Nucl Med* 2003;44:347–52.